matrix algebra for electronic engineers

Paul Hlawiczka

matrix algebra for electronic engineers

Hlawiczka

Iliffe

Although matrix algebra was invented over a century ago, until comparatively recently it remained a subject only of interest to mathematicians. It was then discovered that it provided the neatest, and in some cases the only, method of solving certain problems in physics and electronics. The stage has now been reached when any electronic engineer, or for that matter physicist, without a working knowledge of matrix algebra finds himself severely handicapped.

To the uninitiated, the whole subject looks rather mysterious and most difficult to follow. But nowadays even primary school children are being introduced to it, and certainly any electronic engineer should readily be able to learn sufficient to solve problems involving two-port networks, both active and passive, such as are presented by transistor circuits.

The engineer of today uses mathematics as a tool for solving his problems. Therefore a book, for the engineer, on a particular branch of mathematics should be written by somebody who is fully conversant with both of these fields. The author of the present volume has degrees both in mathematics and engineering; he has been a research engineer with leading electronic companies and has teaching experience. At present he is a Lecturer in Electrical Engineering at Glasgow University.

In this book the subject matter has been chosen to provide a basic course in matrix methods for students of electronic engineering, and for practising engineers who have not had the opportunity of studying the subject before. A pre-view or introduction is included which both smooths the way into a new subject and shows the solution to a number of simple problems in electronics.

The main body of the work is split into two parts. Part 1 is strictly elementary and requires for background only a knowledge of basic algebra. It covers the essentials of matrix methods and their applications to two-port networks, including active networks such as transistor circuits. Part II is distinctly more difficult, but will still be within the capacity of students approaching graduation, and of engineers intending to read papers formulated in terms of matrices. It treats additional algebra up to an explanation of the eigenvalue problem and an introduction of functional concepts relevant to matrices. These methods are then applied to differential equations of linear networks, including the matrix form of the Laplace transformation. The book closes with a chapter on wave matrices, and particularly scattering matrices of microwave junctions, including ferrite devices.

MATRIX ALGEBRA FOR ELECTRONIC ENGINEERS is confidently recommended to students, practising engineers and physicists who wish to master this subject which has now become part of their stock in trade.

216 *pages, including* 71 *diagrams.*

45s net
in UK only

Matrix Algebra for Electronic Engineers

Matrix Algebra for Electronic Engineers

PAUL HLAWICZKA

B.Sc.(Eng), B.Sc., A.M.I.E.E.

Lecturer in Electrical Engineering at the University of Glasgow

LONDON ILIFFE BOOKS LTD

NEW YORK HAYDEN BOOK CO., INC

a division of HAYDEN PUBLISHING COMPANY, INC

First published in 1965 by
Iliffe Books Ltd., Dorset House
Stamford Street, London, S. E. 1

Published in the U. S. A. by
Hayden Book Co., Inc., a division of
Hayden Publishing Company, Inc.,
116 West 14th Street, New York 11

Produced by Uneoprint
set on electric keyboards and with ФОТОМАТН
photo-reproduced and printed offset

at the Gresham Press
UNWIN BROTHERS LIMITED
Old Woking Surrey England

and bound by
THE CHAPEL RIVER PRESS LIMITED
Andover Hampshire England

Contents

Preface

The present book has its origin in a short course of lectures given to an audience of engineers. When expanding the original notes into the form of a book it has been my aim to present the elements of matrix algebra in a form that could be studied profitably by readers without the help of a tutor, and with the minimum of mathematical background. I have, therefore, done everything I could to explain the fundamental concepts fully, and illustrate them copiously.

In the latter part of the book some material beyond the strictly elementary stage is presented. In order to limit the size of the book a selection had to be made. I have chosen topics which are both useful to the engineer and illustrative of the versatility of matrix methods in dealing with problems of some generality. I had in mind the increasing concern which electronic engineers now have for quantum mechanics, by virtue of the emerging new field of quantum electronics. Whatever relevant material there is in this book has been presented in a form adapted to this application. This is particularly the case with the latter sections of Chapter 4 and with Chapter 5.

It is a pleasure to acknowledge my indebtedness to a number of people who influenced this book directly. The original course of lectures was given at the Mid-Essex Technical College, Chelmsford, at the invitation of Mr. D. H. Ray, head of the Electrical Engineering Department. Friends and colleagues of mine who read through parts of the manuscript and offered many constructive criticisms are too numerous to be acknowledged individually. The book is published by permission of the Director of Engineering and Research, The Marconi Company Limited

Glasgow,
1965

P. H.

Part 1

Chapter 1

Introduction

Every electrical engineer is familiar with the concept of a 'black box'. Whether the 'black box' contains an attenuator or a filter section, a vacuum valve or transistor operated linearly, a length of transmission line or a ferrite isolator, its electrical behaviour is completely described by a pair of simple equations.

$$V_1 = AV_2' + BI_2$$
$$I_1 = CV_2' + DI_2 \qquad (1.1)$$

The voltages and currents appearing in these equations are defined in Fig. 1.1. The output voltage V_2' is primed, to stress that its direction is assumed to oppose the current I_2.

The 'black box' is usually referred to as a four terminal network, or a two-port network, and the sum total of its electrical properties is contained in the four coefficients A, B, C, D, —the general circuit constants. Given these constants we can always evaluate the characteristics of the network with the help of Equations 1.1.

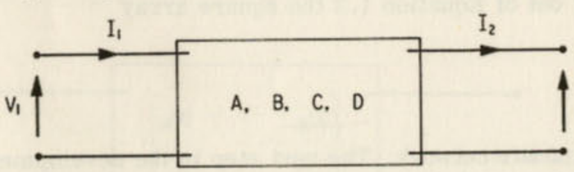


Fig. 1.1

Since 'black boxes' or four terminal networks are frequently encountered in the daily experience of an electrical engineer, it is to be expected that familiarity with their equations will be useful. This is especially the case when several networks have to be handled and interconnected in various ways. Fortunately enough such familiarity is not difficult to acquire since the Equations 1.1 are of the simplest algebraic type. Their outstanding property is their linearity—the electrical variables V_1, I_1, V_2', I_2 appear in the first power, not squared or cubed. In fact, the equations relating currents and voltages in all linear circuits are of this type, so that once their algebra is learned, it should come in useful in many situations of interest to the engineer.

1.1 TERMINOLOGY AND BASIC NOTATION

Equations 1.1 can be considered from two points of view. First, the input voltage and current V_1, I_1 may be known, and it may be desired to find the output voltage and current. In this case Equations 1.1 constitute a set of simultaneous *linear equations* in the unknowns V_2', I_2.

On the other hand, the problem may present itself in reverse. Given specified output quantities V_2', I_2 what voltage and current should be applied to the input? Here we simply substitute the given values of V_2' and I_2 into Equations 1.1 and hence find V_1 and I_1. In this case Equations 1.1 represent a *linear transformation* of the output quantities V_2', I_2 into the input quantities V_1, I_1. On this view a two-port network can be considered to be a device for transforming voltages and currents. The transformer is perhaps the most familiar example of a four terminal network, and its very name suggests this function.

To manipulate linear equations and linear transformations of the form shown in Equations 1.1 a special kind of algebra—*matrix algebra*—has been devised. The first step in the development of this algebra is taken by rewriting Equations 1.1 in the form

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} V_2' \\ I_2 \end{bmatrix} \tag{1.2}$$

Here the arrays of symbols enclosed by square brackets are called *matrices*, and Equation 1.2 is a symbolic way of writing Equations 1.1. We can now present the electrical characteristics of our network in matrix form by isolating out of Equation 1.2 the square array

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

and calling it the matrix of our network. The next step in the development of matrix algebra is to use a single letter to represent the array.

$$A = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \tag{1.3}$$

A is a shorthand symbol standing for the general circuit constants arranged in a square. Likewise the matrices of terminal voltages and currents can be abbreviated by single capital letters in italics.

$$W_1 = \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}, \quad W_2 = \begin{bmatrix} V_2' \\ I_2 \end{bmatrix}$$

Equations 1.1 and 1.2 can now be written in the short symbolic form

$$W_1 = AW_2 \tag{1.4}$$

Matrix algebra amounts to a set of rules, whereby square and rectangular arrays of numbers can be handled through shorthand symbols like Equation 1.4. Once the rules are learned it becomes unnecessary to write out the arrays in full at every stage of a calculation, and much time is saved. Moreover, matrix symbols make it easier to remember complicated relations.

Whenever the word algebra is mentioned operations such as addition, multiplication and division spring to mind. Underlying them all is the concept of equality. In what follows we introduce these concepts and rules for matrices, relying on two-port networks as a guiding example.

1.2 APPLYING MATRIX NOTATION TO SIMPLE CIRCUITS

To start with, consider two distinct four terminal networks as shown in Fig. 1.2.

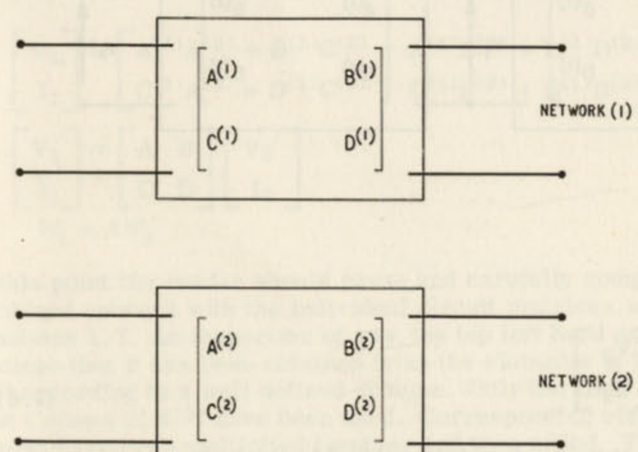


Fig. 1.2

If the two networks have different electrical characteristics their general circuit constants, as marked in Fig. 1.2, will be different complex numbers. If, however, the circuits happen to be identical, their corresponding constants will be equal.

$$A^{(1)} = A^{(2)}, \qquad B^{(1)} = B^{(2)}$$

$$C^{(1)} = C^{(2)}, \qquad D^{(1)} = D^{(2)}$$

In this case the matrices of the networks are said to be equal and we write

$$\begin{bmatrix} A^{(1)} & B^{(1)} \\ C^{(1)} & D^{(1)} \end{bmatrix} = \begin{bmatrix} A^{(2)} & B^{(2)} \\ C^{(2)} & D^{(2)} \end{bmatrix} \tag{1.5}$$

or more briefly

$$A^{(1)} = A^{(2)}$$

Next we turn our attention to the two networks connected together as shown in Fig. 1. 3. We assume that we deal with networks of differing electrical characteristics, having matrices which are not equal. This cascade connection of two networks is itself a four terminal network, and we can write down a pair of equations connecting its input and output voltages and currents, as we did for a single network.

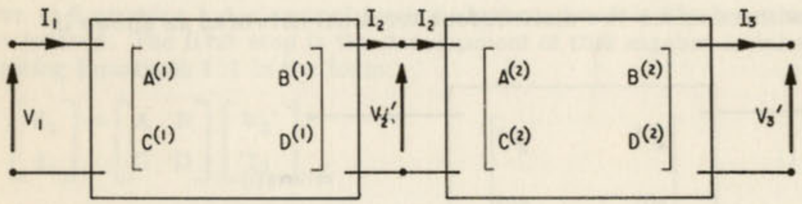


Fig. 1. 3

$$V_1 = AV_3' + BI_3$$
$$I_1 = CV_3' + DI_3$$

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} V_3' \\ I_3 \end{bmatrix} \tag{1.6}$$

The matrix A of the combined network is related to the matrices $A^{(1)}$ and $A^{(2)}$ of the component networks, in the sense that the constants A, B, C, D can be expressed in terms of $A^{(1)}, B^{(1)}, C^{(1)}, D^{(1)}, A^{(2)}, B^{(2)}, C^{(2)}, D^{(2)}$. To find this relation we utilise the fact that the *output* current and voltage of the *first network* are also the *input current* and voltage of the *second network*. First we write down the equations for each individual circuit. (From now on in this introduction we shall write circuit equations in three different forms, to get used to the matrix notation.)

$$V_1 = A^{(1)}V_2' + B^{(1)}I_2, \qquad V_2' = A^{(2)}V_3' + B^{(2)}I_3$$
$$I_1 = C^{(1)}V_2' + D^{(1)}I_2 \qquad I_2 = C^{(2)}V_3' + D^{(2)}I_3$$

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A^{(1)} & B^{(1)} \\ C^{(1)} & D^{(1)} \end{bmatrix} \begin{bmatrix} V_2' \\ I_2 \end{bmatrix}, \qquad \begin{bmatrix} V_2' \\ I_2 \end{bmatrix} = \begin{bmatrix} A^{(2)} & B^{(2)} \\ C^{(2)} & D^{(2)} \end{bmatrix} \begin{bmatrix} V_3' \\ I_3 \end{bmatrix},$$

$$W_1 = A^{(1)} W_2, \qquad W_2 = A^{(2)} W_3 \tag{1.7}$$

Next we substitute the second pair of equations into the first and obtain a relation connecting the input voltage and current V_1, I_1 with the output voltage and current V_3', I_3.

$$V_1 = \left(A^{(1)}A^{(2)} + B^{(1)}C^{(2)}\right)V_3' + \left(A^{(1)}B^{(2)} + B^{(1)}D^{(2)}\right)I_3$$
$$I_1 = \left(C^{(1)}A^{(2)} + D^{(1)}C^{(2)}\right)V_3' + \left(C^{(1)}B^{(2)} + D^{(1)}D^{(2)}\right)I_3$$
$$V_1 = AV_3' + BI_3$$
$$I_1 = CV_3' + DI_3 \tag{1.8}$$

This can be rewritten in matrix form

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A^{(1)}A^{(2)} + B^{(1)}C^{(2)} & A^{(1)}B^{(2)} + B^{(1)}D^{(2)} \\ C^{(1)}A^{(2)} + D^{(1)}C^{(2)} & C^{(1)}B^{(2)} + D^{(1)}D^{(2)} \end{bmatrix} \begin{bmatrix} V_3' \\ I_3' \end{bmatrix}$$

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} V_3' \\ I_3 \end{bmatrix}$$

$$W_1 = AW_3 \tag{1.9}$$

At this point the reader should pause and carefully compare the matrix of the combined network with the individual circuit matrices, as they appear in Equations 1. 7. An inspection of, say, the top left hand corner element will disclose that it has been obtained from the elements of the matrices $A^{(1)}$ and $A^{(2)}$ according to a well defined scheme. Only the first *Row* of $A^{(1)}$ and the first *Column* of $A^{(2)}$ have been used. Corresponding elements of this row and column have been multiplied together and then added. The following graphical scheme will make the procedure clear.

$$\begin{matrix} A^{(1)}B^{(1)} & A^{(2)} \\ & C^{(2)} \end{matrix} \longrightarrow A^{(1)}A^{(2)} + B^{(1)}C^{(2)} = A$$

Similarly the second element of the first row of A has been obtained from the first row of $A^{(1)}$ and the second column of $A^{(2)}$. The remaining elements of A are evaluated by analogous combinations of rows of $A^{(1)}$ and columns of $A^{(2)}$

The above method of combining two matrices to yield a single one is called matrix multiplication. It is written as follows:

$$\begin{bmatrix} A^{(1)} & B^{(1)} \\ C^{(1)} & D^{(1)} \end{bmatrix}\begin{bmatrix} A^{(2)} & B^{(2)} \\ C^{(2)} & D^{(2)} \end{bmatrix} = \begin{bmatrix} A^{(1)}A^{(2)} + B^{(1)}C^{(2)} & A^{(1)}B^{(2)} + B^{(1)}D^{(2)} \\ C^{(1)}A^{(2)} + D^{(1)}C^{(2)} & C^{(1)}B^{(2)} + D^{(1)}D^{(2)} \end{bmatrix}$$

$$A^{(1)}A^{(2)} = A$$

It may well be that at this point the reader will be somewhat bewildered by the variety of symbols used. If so, the difficulty is caused by the symbols themselves, which are not the most efficient to use in connection with matrix manipulations. The symbols devised especially to identify the elements or members of matrix arrays are provided with double subscripts. The first subscript labels the row to which the element belongs, while the second subscript labels the column. Thus, matrices of general circuit constants may be written

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

Taking these arrays for the matrices of the individual networks of Fig. 1.3, the matrix of the combined network will appear in the form

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$$

$$= \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

Using single letter symbols for the matrix arrays we write

$$AB = C$$

Let us now see how these rules are applied to some specific four terminal networks. We start with the series resistance shown in Fig. 1.4.

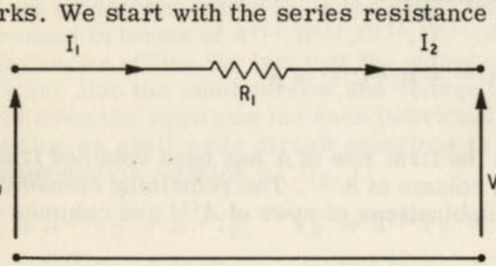


Fig. 1.4

The general circuit constants for this circuit are found by the method of alternately short circuiting and open circuiting the output terminals. On open circuit $I_2 = 0 = I_1$ so that the first of Equations 1.7 assumes the simple form

$$V_1 = A^{(1)}V_2'$$

and since in this case $V_2' = V_1$ we find that

$$A^{(1)} = 1$$

The second equation reduces to

$$I_1 = C^{(1)}V_2' = 0$$

whence

$$C^{(1)} = 0$$

To evaluate the remaining parameters we short circuit the output so that $V_2' = 0$ and $I_2 = I_1 \neq 0$.

Hence we find

$$V_1 = B^{(1)}I_2 = B^{(1)}I_1$$

$$\frac{V_1}{I_1} = B^{(1)} = R_1$$

and for the last constant we obtain

$$I_1 = D^{(1)}I_2 = D^{(1)}I_1$$

$$D^{(1)} = 1$$

Collecting these results and writing them in matrix form, we arrive at the equations of the series resistance connected as a two-port network.

$$\begin{bmatrix} A^{(1)} & B^{(1)} \\ C^{(1)} & D^{(1)} \end{bmatrix} = \begin{bmatrix} 1 & R_1 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} 1 & R_1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} V_2' \\ I_2 \end{bmatrix}$$

$$V_1 = V_2' + R_1 I_2$$

$$I_1 = \quad I_2 \qquad (1.11)$$

Our second example is a resistance in shunt. We label it like the second circuit of Fig. 1.3 with a view to connecting it in cascade with the series element of Fig. 1. 4.

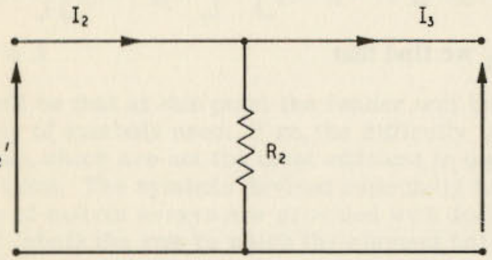


Fig. 1.5

We find the elements of the $A^{(2)}$ matrix for this four terminal network, by the same process of short circuiting and open circuiting its output terminals, as in the first example above. The results are

$$A^{(2)} = \left(\frac{V_2'}{V_3'}\right)_{I_3=0} = 1, \qquad B^{(2)} = \left(\frac{V_2'}{I_3}\right)_{V_3'=0} = 0$$

$$C^{(2)} = \left(\frac{I_2}{V_3'}\right)_{I_3=0} = \frac{1}{R_2}, \qquad D^{(2)} = \left(\frac{I_2}{I_3}\right)_{V_3'=0} = 1$$

Here the subscripts $I_3 = 0$ and $V_3' = 0$ signify that the terminals have been open circuited and short circuited respectively. In matrix notation the equations of the shunt resistance are

$$\begin{bmatrix} V_2' \\ I_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{1}{R_2} & 1 \end{bmatrix} \begin{bmatrix} V_3' \\ I_3 \end{bmatrix}$$

$$\begin{aligned} V_2' &= V_3' \\ I_2 &= \frac{1}{R_2} V_3' + I_3 \end{aligned} \tag{1.12}$$

Let us now connect the series and shunt circuits in cascade, and find the matrix of the resulting L-network by the rule for multiplying matrices. We find

$$\begin{bmatrix} 1 & R_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{1}{R_2} & 1 \end{bmatrix} = \begin{bmatrix} 1 + \frac{R_1}{R_2} & R_1 \\ \frac{1}{R_2} & 1 \end{bmatrix} \tag{1.13}$$

$$A^{(1)} A^{(2)} = L^{(1)}$$

$L^{(1)}$ is the matrix of the circuit shown in Fig. 1.6. The equations relating the input and output of this circuit

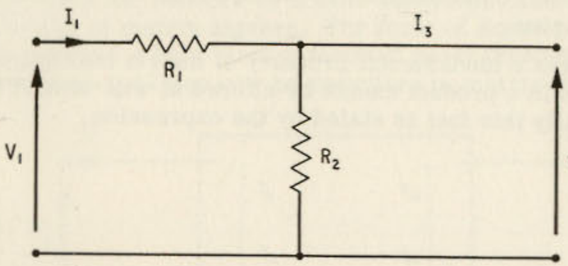


Fig. 1.6

assume the matrix form

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} 1 + \frac{R_1}{R_2} & R_1 \\ \frac{1}{R_2} & 1 \end{bmatrix} \begin{bmatrix} V_3 \\ I_3 \end{bmatrix}$$

$$W_1 = L^{(1)} W_3 \tag{1. 14}$$

or the plain algebraic form

$$\begin{aligned} V_1 &= \left(1 + \frac{R_1}{R_2}\right) V_3' + R_1 I_3 \\ I_1 &= \frac{1}{R_2} V_3' + I_3 \end{aligned}$$

The above result can be obtained by a direct application of Kirchoff's Laws. Indeed, no advantage is gained by the use of matrix multiplication in this simple case, except that it provides a straightforward illustration of the method.

The shunt and series resistances can be connected in reverse order resulting in the circuit of Fig. 1.7. Since the electrical properties of this network differ from those of Fig. 1.6 we must expect its matrix to be different too.

To find its matrix we multiply the matrices of the shunt and series resistances in reverse order to that of Equation 1.13

$$\begin{bmatrix} 1 & 0 \\ \frac{1}{R_2} & 1 \end{bmatrix} \begin{bmatrix} 1 & R_1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & R_1 \\ \frac{1}{R_2} & \frac{R_1}{R_2}+1 \end{bmatrix} \qquad (1.15)$$

$$A^{(2)}A^{(1)} = L^{(2)}$$

The matrix $L^{(2)}$ is *not equal* to the matrix $L^{(1)}$ of Equation 1.13, as has been expected for physical reasons.

This observation illustrates a fundamental property of matrix multiplication. The sequence of matrices in a product cannot be altered at will without affecting the result. Symbolically this fact is stated by the expression

$$A^{(1)}A^{(2)} \neq A^{(2)}A^{(1)}$$

$$L^{(1)} \neq L^{(2)} \qquad (1.16)$$

By contrast we know that ordinary numbers can be multiplied in any order, the final result remaining always the same

$$ab = ba$$

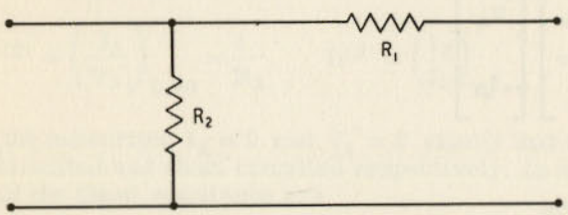


Fig. 1.7

1.3 USE OF ADMITTANCE PARAMETERS

Besides multiplication other rules of algebraic manipulation can be defined for matrices, and many of them resemble symbolically the rules of ordinary algebra. We shall next introduce the addition of matrices on the example of two four terminal networks connected in parallel, instead of in cascade. In this case it proves more convenient to use admittance parameters of a four terminal network, instead of the general circuit constants.

In terms of the admittance parameters the equations of a two-port network assume the form

$$I_1 = y_{11}V_1 + y_{12}V_2$$

$$I_2 = y_{21}V_1 + y_{22}V_2 \qquad (1.17)$$

The output voltage V_2 is now assumed to be in the same direction as the output current, and the prime is therefore dropped. A reference to Fig. 1.8 will make the new sign convention clear. It is usual to label the admittance or y-parameters of the network by double subscripts, automatically conforming to the practice of matrix algebra. The form of Equations 1.17 can be obtained from Equations 1.1 by a straightforward rearrangement of the currents and voltages. Equations 1.17 can now be rewritten in matrix form.

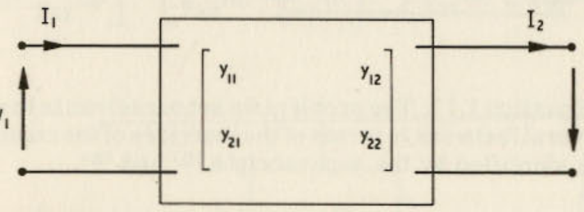


Fig. 1.8

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \qquad (1.18)$$

This can be abbreviated to the compact symbolic equation

$$I = YV \qquad (1.19)$$

Equation 1.19 shows clearly the advantage of matrix symbols in presenting complicated relations in simple form. Just as for a single admittance Y we have the equation

$$I = YV,$$

Equation 1.19 presents the analogous but more involved relation for a four terminal network.

Turning now to the problem of two four terminal networks connected in parallel let us refer to Fig. 1.9.

The corresponding terminals of the circuits are strapped together as shown, resulting in a single two-port network, to be characterised by a pair of

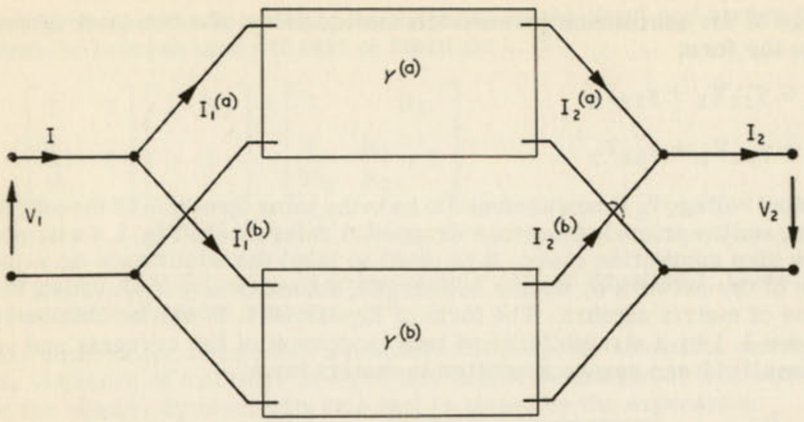


Fig. 1.9

equations of the form of Equation 1.17. The problem we set ourselves is to express the matrix of the overall network in terms of the matrices of the component circuits. The latter are identified by the superscripts $^{(a)}$ and $^{(b)}$.

$$I^{(a)} = Y^{(a)}V$$

$$I_1^{(a)} = y_{11}^{(a)}V_1 + y_{12}^{(a)}V_2$$
$$I_2^{(a)} = y_{21}^{(a)}V_1 + y_{22}^{(a)}V_2 \tag{1.20}$$

$$I^{(b)} = Y^{(b)}V$$

$$I_1^{(b)} = y_{11}^{(b)}V_1 + y_{12}^{(b)}V_2$$
$$I_2^{(b)} = y_{21}^{(b)}V_1 + y_{22}^{(b)}V_2 \tag{1.21}$$

Now, by Kirchhoff's Current Law the currents leaving a junction are equal to the current entering it. Hence we write for the input and output junctions

$$I_1 = I_1^{(a)} + I_1^{(b)}$$
$$I_2 = I_2^{(a)} + I_2^{(b)}$$

Substitution from Equations 1.20 and 1.21 yields the result

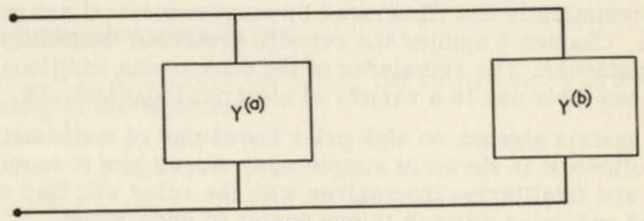
$$I_1 = \left(y_{11}^{(a)} + y_{11}^{(b)}\right)V_1 + \left(y_{12}^{(a)} + y_{12}^{(b)}\right)V_2$$
$$I_2 = \left(y_{21}^{(a)} + y_{21}^{(b)}\right)V_1 + \left(y_{22}^{(a)} + y_{22}^{(b)}\right)V_2 \tag{1.22}$$

Equations 1.22 relate the currents and voltages at the terminals of the overall network, therefore the coefficients in brackets are the y-parameters of the parallel connection. In matrix notation Equation 1.22 can be written

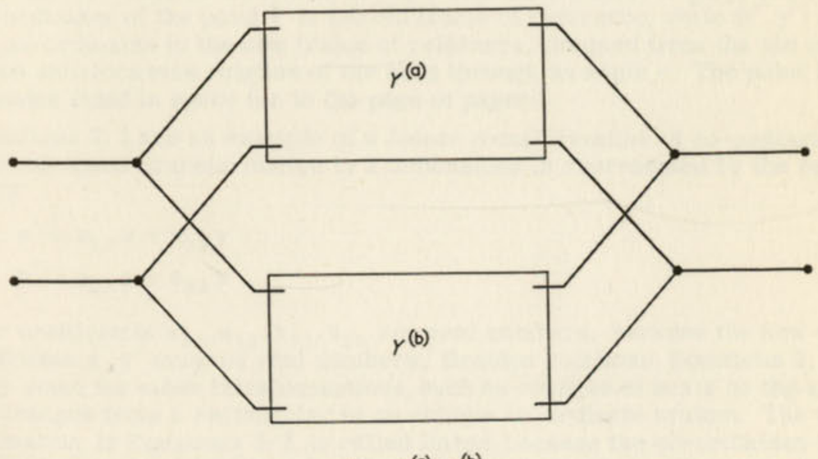
$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} y_{11}^{(a)} + y_{11}^{(b)} & y_{12}^{(a)} + y_{12}^{(b)} \\ y_{21}^{(a)} + y_{21}^{(b)} & y_{22}^{(a)} + y_{22}^{(b)} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \tag{1.23}$$

The matrix of the parallel connection is derived from the matrices of the individual networks by adding corresponding matrix elements. Such a combination of matrices is called addition and is written as follows:

$$\begin{bmatrix} y_{11}^{(a)} & y_{12}^{(a)} \\ y_{21}^{(a)} & y_{22}^{(a)} \end{bmatrix} + \begin{bmatrix} y_{11}^{(b)} & y_{12}^{(b)} \\ y_{21}^{(b)} & y_{22}^{(b)} \end{bmatrix} = \begin{bmatrix} y_{11}^{(a)} + y_{11}^{(b)} & y_{12}^{(a)} + y_{12}^{(b)} \\ y_{21}^{(a)} + y_{21}^{(b)} & y_{22}^{(a)} + y_{22}^{(b)} \end{bmatrix} \tag{1.24}$$



$Y = Y^{(a)} + Y^{(b)}$



$Y = Y^{(a)} + Y^{(b)}$

Fig. 1.10

In terms of one letter symbols this assumes the shorthand form

$$Y^{(a)} + Y^{(b)} = Y$$

The equations of the networks connected in parallel can now be written in the compact form

$$I^{(a)} = Y^{(a)} V, \qquad I^{(b)} = Y^{(b)} V$$

$$I = \left(Y^{(a)} + Y^{(b)}\right) V = YV \tag{1.25}$$

Here again the appeal of the matrix notation will not fail to make its impression. Reference to Fig. 1.10 shows the analogy between ordinary admittances in parallel, and two-port networks in parallel. In the first case the *admittances* are added, in the second case the *admittance matrices* of the two networks are added.

Chapter 1 is a preview of the subjects to be dealt with more thoroughly in the following pages. In Chapter 2 the basic rules of matrix algebra are laid down systematically and illustrated by many numerical and geometrical examples. Chapter 3 applies the rules to problems frequently encountered by circuit engineers. The remainder of the book treats additional algebraic methods and their use in a variety of electrical applications.

To learn matrix algebra no high prior knowledge of mathematics is required. In what follows it is shown in simple, easy stages how to manipulate matrices. Readers who familiarise themselves with the rules will find many problems simplified and many difficult things easier to understand.

Chapter 2

Basic Rules of Matrix Algebra

Historically matrix algebra was invented under the stimulus of problems in analytical geometry. Although it is now possible to develop the subject without any reference to geometrical illustrations, we shall retain the historical connection for two reasons. First it provides helpful pictorial illustrations for the basic manipulations and, indeed, motivates them. Secondly, the geometrical concept most closely related to matrices, linear transformations, helps to clarify many problems in electric circuits. For these reasons we start with a brief discussion of linear transformations, and then lead on to the basic definitions and operations of matrix algebra.

2.1 LINEAR TRANSFORMATIONS

When the axes are rotated about the origin, the co-ordinates of a fixed point change according to the equations

$$\begin{aligned} x' &= x \cos\theta + y \sin\theta \\ y' &= -x \sin\theta + y \cos\theta \end{aligned} \tag{2.1}$$

The symbols used in Equations 2.1 are defined in Fig. 2.1. (x, y) are the co-ordinates of the point P in the old frame of reference, while (x′, y′) are its co-ordinates in the new frame of reference, obtained from the old one by an anticlockwise rotation of the axes through an angle θ. The point P remains fixed in space (or to the page of paper).

Equations 2.1 are an example of a *linear transformation* of co-ordinates. A general linear transformation in 2 dimensions is represented by the equations

$$\begin{aligned} x' &= a_{11} x + a_{12} y \\ y' &= a_{21} x + a_{22} y \end{aligned} \tag{2.2}$$

The coefficients $a_{11}, a_{12}, a_{21}, a_{22}$ are real numbers, because the new co-ordinates x', y' must be real numbers. Besides rotations, Equations 2.2, may stand for other transformations, such as changes of scale on the axes, or changes from a rectangular to an oblique co-ordinate system. The transformation, in Equations 2.2, is called linear, because the co-ordinates or variables under transformation appear to the first power only, and are not squared or cubed.

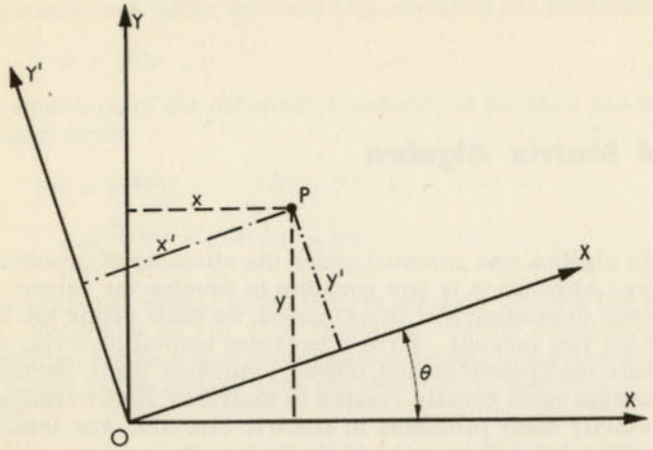


Fig. 2.1

The transformation, in Equations 2.2, can be given an entirely different geometrical interpretation, called *point transformation* instead of co-ordinate transformation. This possibility is illustrated in Fig. 2.2. The frame of reference remains fixed in space (or to the page of paper), but the point $P(x, y)$ is shifted to the position $P'(x', y')$. An alternative way of displaying a point transformation graphically is to use two separate sets of co-ordinate axes, one labelled OXY the other $OX'Y'$ as shown in Fig. 2.3. This method of representation is sometimes called a mapping of the points in the plane OXY on to the plane $OX'Y'$. The law of associating a point P' in the new plane with the point P in the old plane is given by Equations 2.2. The method of mapping is particularly useful when a complexity of points, or possibly loci of points, have to be transformed.

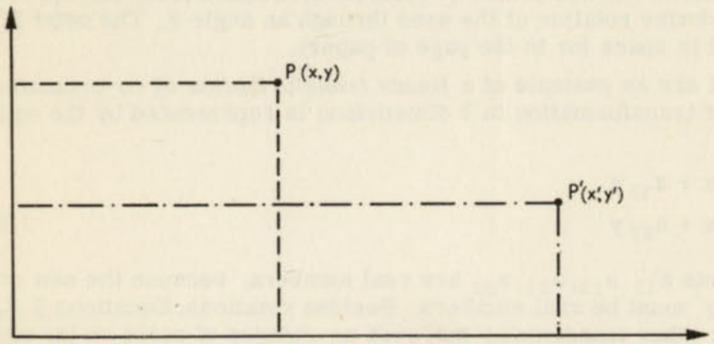


Fig. 2.2

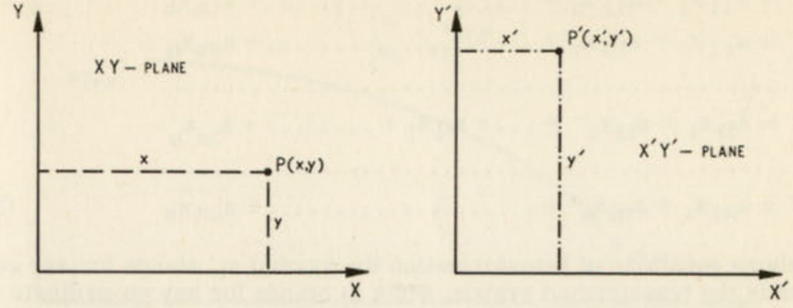


Fig. 2.3

To demonstrate that linear transformations are relevant to electrical problems, we quote the equations relating the input and output of a two-port network. (See Fig. 1.1.)

$$V_1 = AV_2' + BI_2$$
$$I_1 = CV_2' + DI_2$$

The form of these equations is the same as Equations 2.2, A, B, C, D being the coefficients of the transformation. An important difference must be noted, however. Whereas in Equations 2.2 all the symbols stand for real numbers, in electrical relations we have to deal with complex numbers. For this reason transformations involving electrical quantities are more difficult to represent graphically, and we shall restrict ourselves to purely geometrical illustrations.

Linear transformations of co-ordinates or, of points, can be applied to spaces of more than two dimensions. For three dimensional space the equations of transformation are

$$\begin{aligned} x' &= a_{11}x + a_{12}y + a_{13}z \\ y' &= a_{21}x + a_{22}y + a_{23}z \\ z' &= a_{31}x + a_{32}y + a_{33}z \end{aligned} \qquad (2.3)$$

Linear transformations in an abstract n-dimensional space are written down by an extension of the above method. However, in a general case of this type it is impracticable to use a different letter for each co-ordinate. Instead the letter x with a numerical subscript is used. Thus x_1 stands for the first co-ordinate axis, x_2 for the second etc.

$$
\begin{aligned}
x_1' &= a_{11}x_1 + a_{12}x_2 + \ldots\ldots\ldots\ldots\ldots\ldots + a_{1n}x_n \\
x_2' &= a_{21}x_1 + a_{22}x_2 + \ldots\ldots\ldots\ldots\ldots\ldots + a_{2n}x_n \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
x_i' &= a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{ij}x_j + \ldots\ldots + a_{in}x_n \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
x_n' &= a_{n1}x_1 + a_{n2}x_2 + \ldots\ldots\ldots\ldots\ldots\ldots + a_{nn}x_n
\end{aligned}
\tag{2.4}
$$

In the above equations of transformation the symbol x_i' stands for any co-ordinate in the transformed system, while x_j stands for any co-ordinate in the old system. x_n' and x_n stand for the n-th or last co-ordinate. Applying this notation to three dimensions, Equations 2.3 can be rewritten

$$
\begin{aligned}
x_1' &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\
x_2' &= a_{21}x_1 + a_{22}x_1 + a_{22}x_3 \\
x_3' &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3
\end{aligned}
\tag{2.5}
$$

Here x_2' stands for y' and x_3 for z.

The method of labelling the coefficients of a transformation should be noted. The best way to see it is to extract the coefficients from the equations and arrange them in a square array. For Equations 2.2 we find

$$
\begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{matrix}
\tag{2.6}
$$

The first subscript labels the row to which the coefficient belongs, while the second labels the column.

If it is desired to make a statement about any *one of the coefficients* in general, letters are used for subscripts as in a_{ij} of Equations 2.4. On the other hand the symbol a_{ij} is frequently used as an abbreviation for *all the coefficients* of a transformation. The distinction will usually follow from the context, but it is advisable to watch for it initially.

The object of the double subscript notation will become clear in the Section 2.2.

2.2 DEFINITION OF A MATRIX

A linear transformation is fully defined by the set of coefficients appearing in its equations. For example, Fig. 2.4 is a complete description of a linear point transformation, both graphical and algebraic. Given the array of coefficients, equations can be written down immediately, and the co-ordinates of the point P′ evaluated from those of P.

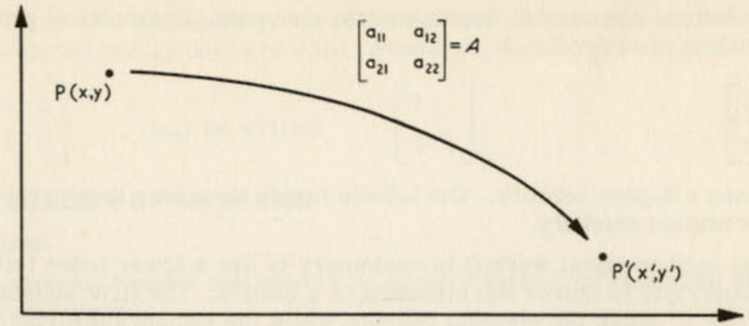


Fig. 2.4

The coefficients $a_{11}, a_{12}, a_{21}, a_{22}$, arranged in a square array, are said to form the *matrix* of the transformation. An important step in the development of a concise notation has been taken by writing them as a neat array, enclosed by square brackets, and equating them to a single symbol *A*. We take the present situation as a basis for the definition of a matrix.

Definition

A rectangular array of numbers consisting of m rows and n columns, enclosed in square brackets, is called a *matrix* of order m × n. The numbers are referred to as *elements* of the matrix.

Some numerical examples of matrices are

order 2 × 2 $\begin{bmatrix} 2 & 3 \\ 1 & 7 \end{bmatrix}$, order 2 × 3 $\begin{bmatrix} 10 & 0 & 3 \\ 5 & 2 & 1 \end{bmatrix}$, order 2 × 2 $\begin{bmatrix} 0{\cdot}2 & 105 \\ 11 & 0{\cdot}35 \end{bmatrix}$

order 3 × 3 $\begin{bmatrix} \tfrac{1}{3} & 2 & 6 \\ 11 & 1 & 0 \\ 0 & 3 & \tfrac{1}{5} \end{bmatrix}$, order 3 × 2 $\begin{bmatrix} \tfrac{1}{9} & \tfrac{3}{16} \\ 2 & \tfrac{3}{8} \\ 0 & 1 \end{bmatrix}$, order 3 × 1 $\begin{bmatrix} 2 \\ 0 \\ 7 \end{bmatrix}$

Matrix elements may be complex numbers as in the following examples:

$$
\begin{bmatrix} 1 + 3j & \tfrac{1}{2} + j \\ 3 & 2 + 6j \end{bmatrix}, \qquad
\begin{bmatrix} 0 & 1 + 2j & 5j \\ 3 + \tfrac{1}{3}j & 0 & 6 + \tfrac{1}{4}j \end{bmatrix}
$$

Sometimes letters are used to denote matrix elements. In electrical problems the matrix of order 2 × 2

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

characterises a 2-port network. The letters inside the array in general represent complex numbers.

For general mathematical work it is customary to use a lower index letter with two subscripts to denote the elements of a matrix. The first subscript labels the row to which the element belongs, while the second subscript labels the column. The whole matrix is usually denoted by an italicised capital letter.

Examples:

$$\begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix} = C, \qquad \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = B$$

A general matrix of order m × n is written as follows:

$$\begin{bmatrix} a_{11} & a_{12} \cdots\cdots\cdots & \cdots a_{1n} \\ a_{21} & a_{22} \cdots\cdots\cdots & \cdots a_{2n} \\ \cdots\cdots & \cdots\cdots\cdots\cdots & \cdots\cdots \\ a_{i1} & a_{i2} \cdots a_{ij} & \cdots\cdots \\ \cdots\cdots & \cdots\cdots\cdots\cdots & \cdots\cdots \\ a_{m1} & a_{m2} \cdots\cdots\cdots & \cdots a_{mn} \end{bmatrix} = A$$

The element a_{ij}, appearing somewhere inside the array, is to be taken as the typical element of the matrix. Other letters may be used as subscripts of the typical element instead of i and j. Thus a_{kl}, a_{pq}, a_{st} will be used as symbols for the typical element in the following pages.

Alternative symbols are sometimes used for matrices in other textbooks. Round brackets or double vertical bars may enclose arrays rather than square brackets. The typical element placed between brackets may be shown instead of a capital italic. Some examples of the various possibilities follow.

$$A, \ \left[a_{ij}\right], \ \left\|a_{ij}\right\|, \ \left(a_{ij}\right)$$

$$\begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix}, \qquad \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix}, \qquad \begin{Vmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{Vmatrix}$$

In the numerical examples given above several zero matrix elements were shown. Occasionally dots are used instead of noughts, as exemplified below.

$$\begin{bmatrix} 1 & 0 \\ 3 & 0 \end{bmatrix} \quad \text{may be written} \quad \begin{bmatrix} 1 & . \\ 3 & . \end{bmatrix}$$

2.3 EQUALITY OF MATRICES

Definition.

Two matrices *A* and *B* are said to be equal if their corresponding elements are equal numbers.

Examples of equal matrices:

$$\begin{bmatrix} 2 & 4 & 7 \\ 9 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 7 \\ 9 & 1 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 1 & e^{j\pi} \\ e^{-j\pi} & 1 \end{bmatrix} = \begin{bmatrix} 1 & e^{j\pi} \\ e^{-j\pi} & 1 \end{bmatrix}$$

On the other hand the following matrices are not equal to each other:

$$\begin{bmatrix} 0 & 3 & 9 \\ 0 & 1 & 0 \end{bmatrix} \neq \begin{bmatrix} 2 & 3 & 9 \\ 6 & 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} a + jb & 1 \\ c + jd & 0 \end{bmatrix} \neq \begin{bmatrix} a & 0 \\ c + jd & 1 \end{bmatrix}$$

Equality of two general matrices, say *A* and *B*, is written in shorthand form

$$A = B$$

or is stated in terms of their typical elements

$$a_{ij} = b_{ij}$$

It should be noted that only matrices having the *same number of rows and columns* can be equal.

$$\begin{bmatrix} 3 & 4 & 7 \\ 9 & 5 & 3 \end{bmatrix} \neq \begin{bmatrix} 3 & 4 & 7 \\ 9 & 5 & 3 \\ 0 & 1 & 6 \end{bmatrix} \neq \begin{bmatrix} 3 & 9 \\ 4 & 5 \\ 7 & 3 \end{bmatrix}$$

When the matrices of two linear transformations are equal, the transformations rotate the co-ordinates by the same angle, if they are rotations. Two point transformations, having equal matrices, shift the point P into the same transformed point P′.

An excellent illustration of the equality of matrices is provided by electric circuits. Referring to Fig. 1.2, the two four terminal networks shown there

have identical electrical characteristics, if their corresponding circuit parameters are equal. In matrix form this means that

$$\begin{bmatrix} A^{(1)} & B^{(1)} \\ C^{(1)} & D^{(1)} \end{bmatrix} = \begin{bmatrix} A^{(2)} & B^{(2)} \\ C^{(2)} & D^{(2)} \end{bmatrix}$$

or more concisely

$$A^{(1)} = A^{(2)}$$

2.4 MULTIPLICATION OF MATRICES

To introduce the product of two matrices we consider the effect of two point transformations applied to the point P in succession. By this we mean the following.

At first we transform the point $P(x_1, x_2)$ into the point $P'(x_1', x_2')$ with the help of the transformation

$$\begin{aligned} x_1' &= a_{11}\,x_1 + a_{12}\,x_2 \\ x_2' &= a_{21}\,x_1 + a_{22}\,x_2 \end{aligned} \tag{2.7}$$

Next we transform the point P' into the point $P''(x_1'', x_2'')$ using the transformation

$$\begin{aligned} x_1'' &= b_{11}\,x_1' + b_{12}\,x_2' \\ x_2'' &= b_{21}\,x_1' + b_{22}\,x_2' \end{aligned} \tag{2.8}$$

Both operations are represented graphically in Fig. 2.5.

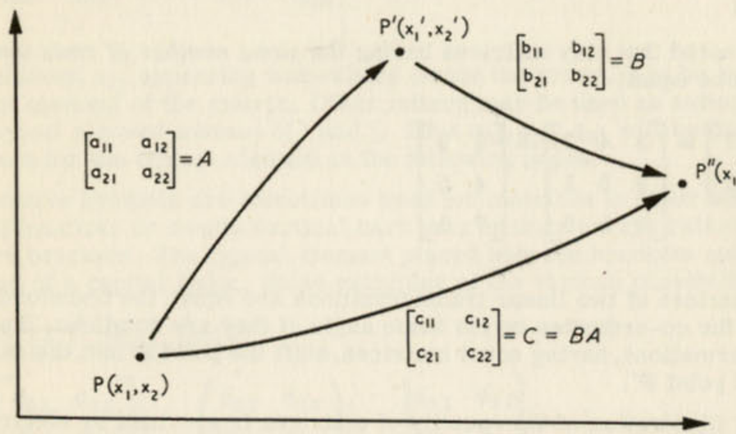


Fig. 2.5

We now substitute Equation 2.7 into Equation 2.8 and obtain the transformation that carries the point P directly into the point P''.

$$\begin{aligned} x_1'' &= \left(b_{11}\,a_{11} + b_{12}\,a_{21}\right)x_1 + \left(b_{11}\,a_{12} + b_{12}\,a_{22}\right)x_2 \\ x_2'' &= \left(b_{21}\,a_{11} + b_{22}\,a_{21}\right)x_1 + \left(b_{21}\,a_{12} + b_{22}\,a_{22}\right)x_2 \end{aligned} \tag{2.9}$$

We can rewrite Equations 2.9 in shorter form by substituting coefficients c_{ij} for the sums of products of the original coefficients a_{ij} and b_{ij}.

$$\begin{aligned} x_1'' &= c_{11}\,x_1 + c_{12}\,x_2 \\ x_2'' &= c_{21}\,x_1 + c_{22}\,x_2 \end{aligned} \tag{2.10}$$

This composite linear transformation replaces the two successive stages.

The point of interest is how to combine the matrices A and B of the separate stages, to obtain the matrix C of the direct transformation. The amalgamation is carried out according to a well defined scheme which we now proceed to explain.

$$\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Consider the elements of the above matrices enclosed in dotted frames. Form the products $b_{11}a_{11}$ and $b_{12}a_{21}$ and add them. The sum is found to be the top left hand corner element of the composite matrix.

$$\begin{bmatrix} b_{11}\,a_{11} + b_{12}\,a_{21} & b_{11}\,a_{11} + b_{12}\,a_{22} \\ b_{21}\,a_{11} + b_{22}\,a_{21} & b_{21}\,a_{12} + b_{22}\,a_{22} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

The element c_{12} is found by the same process applied to the first row of B and the second column of A. In general the subscripts of elements of C indicate which row of B and which column of A was used in their formation *in that order*.

The foregoing method of amalgamating two matrices, by row-into-column multiplication and addition of their elements, is called *matrix multiplication*. Before proceeding any further it is desirable to practise it on some numerical examples.

$$\begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 0 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 22 \\ 4 & 14 \end{bmatrix}$$

The same rules apply to non-square matrices.

$$\begin{bmatrix} 0 & 3 & 1 \\ 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} 5 & 2 \\ 1 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 0+3+0 & 0+0+3 \\ 5+2+0 & 2+0+0 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 7 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2+0 \\ 3+2 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} a+2b \\ c+2d \end{bmatrix}$$

$$\begin{bmatrix} 1 & e^{jx} \\ e^{-jx} & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} e^{jx} & 1 \\ 1 & e^{-jx} \end{bmatrix}$$

Symbolically matrix products are written in the same way as the products of ordinary (or scalar) numbers. Thus

$$BA = C$$

The matrix product

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix}$$

can be used to save time when writing out linear transformations. Using this result Equations 2.7 can be rewritten

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

or more briefly

$$X' = AX$$

In the foregoing examples it was always possible to carry out the multiplication of two matrices. In general this is not the case, however. Two matrices can only be multiplied together, if the first one has as many columns as the second one has rows. The following two matrices cannot form a product.

$$\begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix} \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix}$$

An attempt to carry out the row-into-column multiplication and addition rule, shows that there are no elements in D to match up with c_{13} and c_{23}. This observation is stated in general terms by saying that only matrices of orders $m \times p$ and $p \times n$ can be multiplied.

One more important property of matrix products must be noted. An inspection of the above examples will show that the resulting matrix has as many rows as the first matrix of the product, and as many columns as the second matrix. In general it is said that the product of matrices of orders $m \times p$ and $p \times n$ is a matrix of order $m \times n$.

The rule of matrix multiplication explained above can be summarised in the following definition.

Definition.

The product of two matrices B and A, written BA, is the matrix C whose elements are formed according to the prescription

$$c_{ij} = \sum_{k=1}^{p} b_{ik}a_{kj} \tag{2.11}$$

If the factor matrices B and A are of the orders $m \times p$ and $p \times n$ respectively, the product matrix C is of order $m \times n$.

The use of letter subscripts in Equation 2.11 should be carefully noted. A specific subscript letter, say j, stands for the same number throughout the whole expression. As an example, let us write out Equation 2.11 in full for some elements of the matrix product considered at the beginning of the section.

$$c_{21} = \sum_{k=1}^{2} b_{2k}a_{k1}$$

$$= b_{21}a_{11} + b_{22}a_{21}$$

$$c_{22} = \sum_{k=1}^{2} b_{2k}a_{k2}$$

$$= b_{21}a_{12} + b_{22}a_{22}$$

It is evident that by following the summation rule, Equation 2.11, all elements of a product matrix can be evaluated.

Although many readers may prefer to visualise matrix multiplication in terms of row-into-column manipulations, summations of the form of Equation 2.11 are always used in general mathematical work on matrix products.

2.5 EXAMPLES AND APPLICATIONS OF MATRIX MULTIPLICATION

The rule for multiplying matrices is of fundamental importance. It is, therefore, essential to become quite familiar with it. To this end, let us work through some more numerical examples, this time including negative and complex elements.

$$\begin{bmatrix} 1 & -2 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} -3 & 0 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} -5 & -4 \\ 3 & 5 \end{bmatrix}$$

$$\begin{bmatrix} -1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ -3 & 0 \\ 0 & \tfrac{1}{2} \end{bmatrix}\begin{bmatrix} -2 & 0 & 1 \\ 0 & -\tfrac{1}{2} & 3 \end{bmatrix} = \begin{bmatrix} -2 & -1 & 7 \\ 6 & 0 & -3 \\ 0 & -\tfrac{1}{4} & \tfrac{3}{2} \end{bmatrix}$$

$$\begin{bmatrix} 1+j & 3-2j \\ 0 & -j \end{bmatrix}\begin{bmatrix} 0 & \tfrac{1}{4}+\tfrac{1}{2}j \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3-2j & \tfrac{1}{4}(23-13j) \\ -j & -2j \end{bmatrix}$$

$$\begin{bmatrix} r_1e^{j\alpha} & 1 \\ 1 & r_2e^{j\beta} \end{bmatrix}\begin{bmatrix} ae^{j\gamma} \\ be^{j\delta} \end{bmatrix} = \begin{bmatrix} ar_1e^{j(\alpha+\gamma)} + be^{j\delta} \\ ae^{j\gamma} + br_2e^{j(\beta+\delta)} \end{bmatrix}$$

$$\begin{bmatrix} a+jb & c+jd \\ d+je & jf \end{bmatrix}\begin{bmatrix} a-jd \\ -jg \end{bmatrix} = \begin{bmatrix} (a^2+bd+dg)+j(ba-ad-cg) \\ (ad-de+fg)-j(ae+d^2) \end{bmatrix}$$

It was pointed out in the preceding section that the equations of a linear transformation can be written in a concise form using matrix symbols.

$$X' = AX \qquad (2.12)$$

Equation 2.12 represents a transformation in any number of dimensions, not necessarily in the plane. The information about the number of dimensions involved in a linear transformation is thus lost, when matrix notation is used and, if necessary, must be provided by the context. This is the price that must be paid when a more abstract mathematical method is used.

It was pointed out in Section 2.1 that the equations relating the input and output of a two-port network are a linear transformation. With the help of the

rule for multiplying matrices we can now rewrite these equations in compact form.

$$\begin{aligned} V_1 &= AV_2' + BI_2, \\ I_1 &= CV_2' + DI_2 \end{aligned} \qquad \begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} V_2' \\ I_2 \end{bmatrix}$$

$$W_1 = AW_2$$

Matrices consisting of a single row or column of elements are usually called *row* and *column vectors* respectively. As further examples of matrix products let us multiply a row vector by a column vector and vice versa.

$$XY = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_1y_1 + x_2y_2 + x_3y_3 \end{bmatrix}$$

$$YX = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} = \begin{bmatrix} y_1x_1 & y_1x_2 & y_1x_3 \\ y_2x_1 & y_2x_2 & y_2x_3 \\ y_3x_1 & y_3x_2 & y_3x_3 \end{bmatrix}$$

In the second case a matrix of order 3×1 has been multiplied by one of order 1×3 resulting in a matrix of order 3×3. In the first case we started with matrices of orders 1×3 and 3×1 yielding a product matrix of order 1×1. It is useful to note that matrices of order 1×1 have the same mathematical properties as ordinary (or scalar) numbers. This will be demonstrated in detail later on, but for the time being we accept it that a 1×1 matrix is the same as a number.

The rules of matrix multiplication can be applied to write down systems of simultaneous linear equations in concise form. Thus in the case of three linear equations in three unknowns x, y, z, we can write

$$\begin{aligned} a_{11}x + a_{12}y + a_{13}z &= h_1, \\ a_{21}x + a_{22}y + a_{23}z &= h_2 \\ a_{31}x + a_{32}y + a_{33}z &= h_3 \end{aligned} \qquad \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{23} & a_{33} \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} \begin{matrix} = \\ = \\ = \end{matrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}$$

In abbreviated form this is

$$AX = H \qquad (2.13)$$

where A is the matrix of coefficients, X stands for the column vector of unknowns, and H is the column vector of constants on the right hand side of the set of equations.

Equation 2.13 shows the degree of compression achieved by using matrix symbols to write down systems of equations. However, the advantages of

matrix methods go far beyond affording a mathematical shorthand. We shall see later that matrices will provide an elegant method for solving equations. In fact, we shall anticipate our future work to some extent now. Just as we solve a single equation in one unknown

$$ax = h$$

by transferring a to the right hand side

$$x = a^{-1}h$$

we shall be able to solve a whole system of equations by using the reciprocal matrix A^{-1}.

$$X = A^{-1} H$$

Linear equations do not always appear in simple forms like the above. To begin with, the number of equations and unknowns may not be the same. Matrix symbols and matrix multiplication are still applicable. In the case of two equations in three unknowns we write

$$\begin{aligned} a_{11}x + a_{12}y + a_{13}z &= h_1, \\ a_{21}x + a_{22}y + a_{23}z &= h_2 \end{aligned} \qquad \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}$$

$$AX = H$$

The same abbreviation as in Equation 2.13 applies here.

Sometimes all the constants on the right hand side of Equation 2.13 may be zero. The equations then assume the form

$$\begin{aligned} a_{11}x + a_{12}y + a_{13}z &= 0, \\ a_{21}x + a_{22}y + a_{23}z &= 0 \\ a_{31}x + a_{32}y + a_{33}z &= 0 \end{aligned} \qquad \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{23} & a_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad (2.14)$$

$$AX = 0$$

On the extreme right we have a rather special type of matrix called the *null matrix*. This name is given to matrices whose elements are all zeros. If not written out in full, a null matrix can be denoted by any one of the symbols

$$A = O, \quad \left[a_{ij}\right] = \left[0\right] = 0, \quad a_{ij} = 0 \text{ for all } i \text{ and } j.$$

The following examples should convince the reader that multiplication of any matrix by a null matrix always results in a null matrix.

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ -2 & 0 \\ 6 & 11 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Hence, a null matrix is analogous to the zero among scalar numbers.

An interesting property of matrix products may be noted at this point. Although the product of two matrices may vanish, it does not follow that one of the factors must be a null matrix. Thus, even if we know that

$$AB = 0$$

we cannot conclude that either $A = 0$ or $B = 0$. To verify this consider the example

$$\begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Hence it is quite possible that the product AX may be a null matrix without either A or X being a null matrix. This important observation is connected with the known fact that a system of equations as in Equations 2.14 may have a non-zero solution.

Further examples of matrix multiplication may be found at the end of this chapter. The reader is urged to practise on them until a confident grasp of the mechanism of row-into-column multiplication is acquired.

2.6 RULES OF MATRIX MULTIPLICATION

There is a symbolic resemblance between products of matrices and products of scalar numbers. Given two numbers a and b, their product is written ab, which appears similar to the product of matrices A and B

$$AB$$

This symbolic similarity must not be allowed to suggest that matrix products obey the same rules as products of numbers. Whereas two numbers can be multiplied in any sequence without affecting the final result, this is not so with matrices. In symbols the fact is stated as follows:

$$ab = ba \qquad (2.15)$$

$$AB \neq BA \qquad (2.16)$$

A demonstration of this rule is provided by a numerical example picked almost at random.

$$AB = \begin{bmatrix} 0 & -1 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 4 & -2 \end{bmatrix} = \begin{bmatrix} -4 & 2 \\ 16 & -6 \end{bmatrix} = C$$

$$BA = \begin{bmatrix} 2 & 0 \\ 4 & -2 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 0 & -2 \\ -4 & -10 \end{bmatrix} = D$$

The product matrices C and D are clearly not equal, so that the example verifies the *non-commutative law of matrix multiplication.*

Apart from the foregoing example the non-commutative law can be seen to operate in the case of products of rectangular matrices. Take, for example, the product of a row and a column vector. If the row vector is placed first, we obtain as a result a matrix of order 1×1. If the order of factors is reversed, we find a product matrix of order $n \times n$ (n = number of elements in each vector). By definition these results are not equal.

In some cases the reversed multiplication cannot be carried out at all. Thus, if we are given two rectangular matrices A and B (of orders $p \times n$ and $n \times q$ respectively, the product AB exists and is of order $p \times q$, whereas BA cannot be formed at all if $p \neq q$.

The non-commutative law merely states that matrix products do not always commute. It allows the possibility that some matrices may, in fact, commute under multiplication. We shall meet examples of such matrices as we go on. In view of the non-commutative law, it is sometimes necessary to distinguish between the factors in a matrix product. This is accomplished by the terms *premultiplication* and *postmultiplication*. We say that B is premultiplied by A in the expression AB, but it is postmultiplied in BA.

Linear transformations give a geometrical meaning to the non-commutative law. Consider point transformations in the plane, starting with the point P, as shown in Fig. 2.6.

The product transformation $BA = C$ carries the point P into P_{BA}, whereas $AB = D$ carries it into P_{AB}. An inspection of the equations of transformation will make it clear that in general the points P_{BA} and P_{AB} do not coincide, just as the products C and D are not necessarily equal.

So far we have considered products of two matrices only. A moments reflection will show that matrix multiplication can be extended to include any number of factors. Thus we can multiply two matrices, say A and B, to form the product $AB = E$, and then to postmultiply this by C, giving

$$EC = (AB)C = P \qquad (2.17)$$

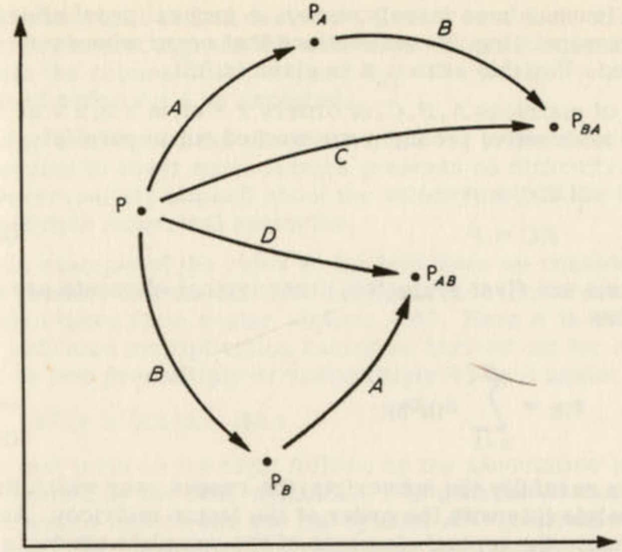


Fig. 2.6

Evidently the procedure can be continued up to any number of matrices.

An important question arises regarding such multiple products. For the purpose of calculating the product shown in Equation 2.17 the matrices A and B were first grouped together and then postmultiplied by C. It would be possible to group B and C together and then premultiply them by A.

$$A(BC)$$

Would the final result be the same as in Equation 2.17? The answer is yes. In a multiple product of matrices the factors may be grouped in any way, without affecting the final result. Hence we can write

$$(AB)C = A(BC) = ABC = P \qquad (2.18)$$

By the non-commutative law of matrix multiplication the *sequence* of factors must, however, be carefully preserved.

As regards the grouping of factors, matrix products are symbolically similar to products of numbers. Three numbers, say a, b, c can be multiplied according to either grouping.

$$(ab)c = a(bc) = p$$

The above property of products, common to both numbers and matrices, is called the *associative law of multiplication.*

So far the associative law has been merely stated. A general proof affords excellent practice in manipulating the summations that occur whenever matrices are multiplied. For this reason it is given in full.

We consider products of matrices A, B, C, of orders $r \times m$, $m \times n$, $n \times s$, respectively. The two alternative products are worked out in parallel:

$$\begin{aligned} A(BC) &= P' \qquad & (AB)C &= P \\ AD &= P' \qquad & EC &= P \end{aligned} \tag{2.19}$$

The products in brackets are first evaluated; their typical elements are given by the summations

$$d_{hj} = \sum_{k=1}^{n} b_{hk}c_{kj}, \qquad e_{ik} = \sum_{h=1}^{m} a_{ih}b_{hk} \tag{2.20}$$

The reader should note carefully the subscripts, the ranges over which they are summed, and correlate this with the order of the factor matrices. As the next step in the calculation the typical elements of the complete products are written down.

$$p_{ij}' = \sum_{h=1}^{m} a_{ih}d_{hj}, \qquad p_{ij} = \sum_{k=1}^{n} e_{ik}c_{kj} \tag{2.21}$$

Equations 2.20 are now substituted in Equations 2.21:

$$p_{ij}' = \sum_{h=1}^{m} a_{ih}\left(\sum_{k=1}^{n} b_{hk}c_{kj}\right), \qquad p_{ij} = \sum_{k=1}^{n}\left(\sum_{h=1}^{m} a_{ih}b_{hk}\right)c_{kj} \tag{2.22}$$

Here we have to do with finite sums of ordinary numbers, hence the double summations can be carried out in any order:

$$p_{ij}' = \sum_{h=1}^{m}\sum_{k=1}^{n} a_{ih}b_{hk}c_{kj}, \qquad p_{ij} = \sum_{h=1}^{m}\sum_{k=1}^{n} a_{ih}b_{hk}c_{kj} \tag{2.23}$$

This leaves us with the conclusion

$$p_{ij}' = p_{ij} \tag{2.24}$$

The corresponding elements of the product matrices are equal and we have proved the desired result

$$P' = P \tag{2.25}$$

The subscript i ranges over the values 1 to r since it is the row label of the matrix element a_{ih}, and the subscript j has the range from 1 to s since it labels the column of the element c_{kj}. The product matrix P or P' is, therefore, of order $r \times s$ as expected.

Having established the associative law for products of three matrices, its extension to cover more factors presents no difficulty. The reader should, however, satisfy himself about the validity of this law by working out in full a few simple numerical examples.

As an example of the rules so far laid down we consider powers of matrices. The product of a matrix with itself, say AA, can be denoted by the squaring symbol taken from scalar algebra - A^2. Here A is necessarily square since the indicated multiplication cannot be carried out for a non-square matrix. Let us now premultiply or postmultiply A^2 by A again:

$$(A^2)A = (AA)A = AAA$$

The last term on the right follows by the associative law and it is conveniently denoted by the cube symbol A^3. In general we can form any integral power of a square matrix and use the symbol A^n. It should be observed that negative or fractional powers of matrices have not been defined so far.

Before closing the sections on matrix multiplication we note in retrospect that this operation is symbolically similar to the multiplication of scalar numbers. In fact, both the terminology and symbols are all drawn from ordinary algebra. The one notable exception is the non-commutative property of matrix products. It will be borne in mind, however, that the procedure for multiplying actual matrices is relatively complicated, as it involves both addition and multiplication of ordinary numbers.

2.7 MULTIPLICATION OF A MATRIX BY A NUMBER

Definition

A matrix is said to be multiplied by a scalar number if all its elements are multiplied by this number. In symbols:

$$kA = \left[ka_{ij}\right]$$

Examples:

$$2\begin{bmatrix} 3 & 7 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 14 \\ 0 & 2 \end{bmatrix}$$

$$2\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 & 2x_2 & 2x_3 \end{bmatrix}$$

$$k\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} = \begin{bmatrix} ka & kb & kc \\ kd & ke & kf \end{bmatrix}$$

There are both algebraic and geometrical reasons why this particular definition should be adopted, they will become clear as the subject develops. For the present the reader is warned not to confuse this type of matrix multiplication with a well known rule for determinants. The distinction will be brought out more fully later.

It is sometimes found that a scalar factor is involved in a multiple product such as AkB. It is then permissible to take the scalar outside the product of matrices and use it only in the last stage of calculation.

Thus

$$AkB = kAB \tag{2.26}$$

An example will help to explain why this should be so

$$AkB = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} kb_{11} & kb_{12} \\ kb_{21} & kb_{22} \end{bmatrix}$$

$$= \begin{bmatrix} k(a_{11}\,b_{11} + a_{12}\,b_{21}) & k(a_{11}\,b_{12} + a_{12}\,b_{22} \\ k(a_{21}\,b_{11} + a_{22}\,b_{21}) & k(a_{21}\,b_{11} + a_{22}\,b_{22} \end{bmatrix}$$

$$= kAB$$

It will be observed that k appears as a linear factor in all the products of numbers making up the elements of the product matrix. Hence it can be taken outside the brackets, and in the end extracted outside the matrix product.

The above rule extends to products of any number of matrices and scalars. In the general case the scalar factors can be shifted between the matrices at will.

e.g.

$$A\,kB\,lC = kl\,A\,B\,C = A\,B\,kl\,C \text{ etc.}$$

2.8 ADDITION OF MATRICES

Matrix addition is so defined that only matrices of the same order can be added. The prescription for adding them is given by the following definition.

Definition

Addition of two matrices, both of order $m \times n$, is accomplished by adding their corresponding elements. As a result a matrix of order $m \times n$ is obtained, given by the equation

$$A + B = \left[a_{ij}\right] + \left[b_{ij}\right] = \left[a_{ij} + b_{ij}\right] = \left[c_{ij}\right] = C \tag{2.27}$$

Examples:

$$\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} + \begin{bmatrix} g & h \\ k & l \\ p & q \end{bmatrix} = \begin{bmatrix} a+g & b+h \\ c+k & d+l \\ e+p & f+q \end{bmatrix}$$

$$[0 \;\; 1 \;\; 3] + [-2 \;\; 0 \;\; 5] = [-2 \;\; 1 \;\; 8]$$

An instructive example of matrix addition is provided by a matrix with complex elements or, briefly, a *complex matrix*. Just as a complex number is the sum of two real numbers, one of them multiplied by the imaginary unit j, so a complex matrix is the sum of two real matrices, one of them multiplied by j.

$$\begin{bmatrix} a+jb & c+jd \\ e+jf & g+jh \end{bmatrix} = \begin{bmatrix} a & c \\ e & g \end{bmatrix} + \begin{bmatrix} jb & jd \\ jf & jh \end{bmatrix}$$

$$= \begin{bmatrix} a & c \\ e & g \end{bmatrix} + j\begin{bmatrix} b & d \\ f & h \end{bmatrix}$$

In the last line the rule for the multiplication of a matrix by a scalar number has been applied.

The addition rule for matrices can be applied to two four terminal networks connected in parallel. The admittance matrices of the networks are used for the purpose. In the notation of Fig. 1.9 we have

$$Y^{(a)} + Y^{(b)} = Y$$

In full we write

$$\begin{bmatrix} y_{11}^{(a)} & y_{12}^{(a)} \\ y_{21}^{(a)} & y_{22}^{(a)} \end{bmatrix} + \begin{bmatrix} y_{11}^{(b)} & y_{12}^{(b)} \\ y_{21}^{(b)} & y_{22}^{(b)} \end{bmatrix} = \begin{bmatrix} y_{11}^{(a)} + y_{11}^{(b)} & y_{12}^{(a)} + y_{12}^{(b)} \\ y_{21}^{(a)} + y_{21}^{(b)} & y_{22}^{(a)} + y_{22}^{(b)} \end{bmatrix}$$

Matrix addition, like the addition of numbers, is both commutative and associative. That is, we can always write

$$A + B = B + A \tag{2.28}$$

$$A + (B + C) = (A + B) + C \tag{2.29}$$

The truth of these statements is obvious once it is realised that only the addition of numbers is really involved. To make sure that the Equations 2.28 and 2.29 are clearly understood, the examples above should be rewritten with the addends placed in reversed order.

2.9 DISTRIBUTIVE LAW

An algebraic rule for matrices will now be established that involves both addition and multiplication. The *distributive law* for matrices assumes the same sumbolic form as for scalars. In the latter case we can always write

$$(a + b)c = ac + bc$$

For matrices the same relation holds.

$$(A + B)C = AC + BC \tag{2.30}$$

A proof of the distributive law can be obtained by writing down in full the summations that yield typical elements of the matrices on both sides of Equation 2.30. It is then seen that they are equal. This is left as an exercise for the reader.

In connection with the distributive law it must be remembered that matrix multiplication is non-commutative.

Thus in the product

$$C(A + B) = CA + CB$$

the factor C must precede A and B just as in Equation 2.30 it must follow them.

2.10 REMARKS AND EXAMPLES

With the exception of division all the basic manipulations of matrix algebra have now been explained. It is desirable to pause for a while before introducing this last operation, to take stock and to consider some more examples.

Symbolically matrix algebra is very much like the algebra of scalars, with the exception of the non-commutative law of multiplication. To bring out this fact we summarise the rules for matrices and numbers side by side, in Table 2.1.

In view of these laws we are allowed to manipulate matrix expressions in a way similar to scalars, paying due attention to the sequence of factors in any matrix product. Thus, to form the cube of a sum of matrices we proceed as follows:

$$(A + B)^3 = A^3 + A^2B + AB^2 + BA^2 + B^2A + BAB + ABA + B^3$$

The terms on the right hand side must be left separate. Because of the non-commutative law, they cannot be grouped into four terms as in the analogous relation for scalars. Next consider the expression

$$(A + B)\,(A - B) = A^2 + BA - AB - B^2$$

Table 2.1.

Matrices.	Scalar numbers.
Commutative law of addition. $A + B = B + A$	$a + b = b + a$
Associative law of addition. $A + (B + C) = (A + B) + C$	$a + (b + c) = (a + b) + c$
Non-commutative law of matrix multiplication. $AB \neq BA$	Commutative law of scalar multiplication. $ab = ba$
Associative law of multiplication. $(AB)C = A(BC)$	$(ab)c = a(bc)$
Distributive law. $A(B + C) = AB + AC$ $\neq BA + CA$	$a(b + c) = ab + ac$ $= ba + ca$

This should be contrasted with the factored form of the difference of squares of scalar numbers.

$$(a + b)\,(a - b) = a^2 - b^2$$

Many more such examples can be written down and compared with similar relations for scalars.

2.11 SPECIAL TYPES OF MATRICES

The null matrix introduced in Section 2.5 had a special property in that all its elements were zeros. In the present section we introduce further examples of special matrices, starting with the unit matrix. This is a square matrix which has unit elements in the principal diagonal and zeros everywhere else. The unit matrix of order 3×3 is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The *principal diagonal* of a matrix, or briefly the diagonal, starts in the top left hand corner. The remaining elements of a matrix array are frequently referred to as off diagonal.

The outstanding property of the unit matrix is that it leaves any matrix unaltered under multiplication. Therefore it is analogous to unity among scalars.

Example:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

In general we can write for the unit matrix

$$AI = IA = A \tag{2.31}$$

where I stands for the unit matrix.

It should be noted that unit matrices used in pre- or postmultiplication with a rectangular matrix will not be of the same order. Thus, in the example above, the unit matrix must be of order 2×2 if used as a prefactor:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

The unit matrix, or rather its elements, can be described by the equation

$$a_{ij} = \delta_{ij} \tag{2.32}$$

where δ_{ij} is the *Kronecker*-δ symbol. It is defined by the algebraic property

$$\delta_{ij} = \begin{cases} 1 \text{ when } i = j \\ 0 \text{ when } i \neq j \end{cases} \tag{2.33}$$

In connection with linear transformations the unit matrix represents the *identical transformation.*

$$X = IX = X \tag{2.34}$$

In two dimensions this is

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Considered as a point transformation the identical transformation leaves the point $P(x_1, x_2)$ unmoved relative to the frame of reference.

Next consider a square matrix of the special form

$$\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$

As in the unit matrix the off diagonal elements are zero, but the diagonal elements can be any numbers. This matrix is called *diagonal*. It can be briefly described using the Kronecker-δ symbol.

$$\left[a_{ij}\right] = \left[a_{ij}\delta_{ij}\right] \tag{2.35}$$

Diagonal matrices provide an exception to the non-commutative law of multiplication.

Example:

$$\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \begin{bmatrix} d & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & f \end{bmatrix} = \begin{bmatrix} ad & 0 & 0 \\ 0 & be & 0 \\ 0 & 0 & cf \end{bmatrix}$$

$$\begin{bmatrix} d & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & f \end{bmatrix} \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} = \begin{bmatrix} da & 0 & 0 \\ 0 & eb & 0 \\ 0 & 0 & fc \end{bmatrix}$$

A general proof that diagonal matrices commute under multiplication provides excellent practice in the use of the important Kronecker-δ symbol.

We start with two diagonal matrices $D^{(1)}$ and $D^{(2)}$, both of order $n \times n$, and we set out to show that

$$D^{(1)}D^{(2)} = D^{(2)}D^{(1)} = D \tag{2.36}$$

Starting with the left hand side the typical element of the product is

$$d_{ij} = \sum_{k=1}^{n} d_{ik}{}^{(1)} d_{kj}{}^{(2)}$$

$$= \sum_{k=1}^{n} \left(d_{ik}{}^{(1)}\delta_{ik}\right)\left(d_{kj}{}^{(2)}\delta_{kj}\right)$$

with the help of Equation 2.35. The first term in brackets under the summation sign vanishes, except for $k = i$, by the property of the Kronecker symbol.

Hence the summation reduces to a single term for k = i and we find

$$d_{ij} = \left(d_{ii}^{(1)}\delta_{ii}\right)\left(d_{ij}^{(2)}\delta_{ij}\right)$$

We simplify this by virtue of $\delta_{ii} = 1$ to the form

$$d_{ij} = d_{ii}^{(1)}\left(d_{ij}^{(2)}\delta_{ij}\right)$$

Here we use the property of the Kronecker symbol again and conclude that the above product vanishes except for j = i. Hence the only non-vanishing elements of the product $D^{(1)}D^{(2)}$ are the diagonal elements

$$d_{ii} = d_{ii}^{(1)}d_{ii}^{(2)} \tag{2.37}$$

An argument like the above for the product $D^{(2)}D^{(1)}$ leads again to Equation 2.37. Hence diagonal matrices can be multiplied in any order as stated by Equation 2.36. Moreover, according to Equation 2.37, diagonal matrices are multiplied by multiplying together their corresponding diagonal elements.

It is sometimes convenient to include rectangular matrices under the heading diagonal. Thus a matrix of the form

$$\begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \end{bmatrix}$$

may be referred to as diagonal, although it is not strictly within the above definition. Rectangular diagonal matrices do not, of course, commute under multiplication.

An important type of special matrix is obtained by letting the non-zero elements of a diagonal matrix be equal numbers

Example of order 3 × 3:

$$\begin{bmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & k \end{bmatrix}$$

The name given to this type is *scalar matrix*. The reason for this name becomes apparent when any matrix is multiplied by a scalar matrix.

$$\begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} = \begin{bmatrix} ka_{11} & ka_{12} & ka_{13} \\ ka_{21} & ka_{22} & ka_{23} \end{bmatrix}$$

The effect of multiplying any matrix by a scalar matrix is the same as multiplication by a scalar number.

In terms of the Kronecker-δ symbol the scalar matrix can be written

$$\left[a_{ij}\right] = \left[k\delta_{ij}\right]$$

It can also be expressed in terms of the unit matrix, when it is observed that the scalar matrix can be obtained from the unit on multiplication by the appropriate scalar.

$$kI = \left[k\delta_{ij}\right]$$

The properties of the scalar matrix are summarised in the relation

$$\left[kb_{ij}\right] = kB = kIB \tag{2.38}$$

where B is a general matrix.

Inserting −1 for the diagonal elements of a scalar matrix we obtain a matrix of the form

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = -1\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = -1\,I = -I$$

The above matrix is clearly analogous to −1 among numbers. Its effect on any matrix A is the same as multiplication by the scalar −1, and the operation is denoted by any one of the following symbols:

$$-I\,A = -1\,A = -A$$

Thus changing the sign of a matrix is equivalent to multiplying it by $-I$ or −1.

A matrix is called *symmetric* if the elements situated on opposite sides of the principal diagonal are equal.

Examples:

$$\begin{bmatrix} 1 & 4 & 5 \\ 4 & 2 & 6 \\ 5 & 6 & 3 \end{bmatrix}, \quad \begin{bmatrix} a & d & e \\ d & b & f \\ e & f & c \end{bmatrix}$$

When we come to define the reciprocal of a matrix in a subsequent section, we shall find it necessary to rearrange a matrix array by interchanging its rows and columns. This operation is so important that it is given a special name: *transposition*. A matrix with interchanged rows and columns is called the *transpose* of the original matrix.

Examples:

original matrix $\qquad$ its transpose

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \qquad \begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}$$

The transpose of a matrix A is usually denoted by a prime: A'. Sometimes the symbol A^t is also used. Another method of denoting the transposed matrix is to interchange the subscripts of its elements

$$A' = \left[a_{ij}\right]' = \left[a_{ji}\right].$$

Transposition provides a method of distinguishing between row and column vectors. Thus, if X represents a column vector, X' represents the corresponding row vector. To avoid confusion between the transpose of a column vector and the coordinates in a linear transformation, the latter will henceforth be denoted by superscripts in brackets, e.g. $X^{(1)}, X^{(2)}$ etc.

The symmetric property of a matrix may be conveniently stated in terms of transposition. Thus, if S is any symmetric matrix, it satisfies the relation

$$S' = S \tag{2.39}$$

In terms of typical elements Equation 2.39 can be written

$$\left[s_{ij}\right] = \left[s_{ji}\right] \tag{2.40}$$

or

$$s_{ij} = s_{ji}$$

It should be noted that only square matrices can be symmetric.

Symmetric matrices occur frequently in the equations of linear, passive, bilateral electric circuits, as will be seen in the next chapter.

Let us now consider the following matrices:

$$\begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ 2 & -3 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & -a-jb & -c+jd \\ a+jb & 0 & -e-jf \\ c-jd & e+jf & 0 \end{bmatrix}$$

The elements on opposite sides of the principal diagonal are numerically equal, but have opposite signs. Take one of these matrices, transpose it, and change its sign (or multiply it by -1 or $-I$). Denoting the matrix by the symbol A, the sequence of operations is

$$A = \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{bmatrix}, \quad A' = \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -3 \\ 2 & 3 & 0 \end{bmatrix}, \quad -A' = \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{bmatrix}$$

The matrix A is seen to satisfy the relation

$$-A' = A \quad \text{or} \quad A' = -A \tag{2.41}$$

Matrices of this type are said to be *antisymmetric* or *skew symmetric*. Statements of the antisymmetric property of a matrix, alternative of Equations 2.41 are

$$\left[a_{ij}\right] = \left[-a_{ji}\right], \quad a_{ij} = -a_{ji}$$

The diagonal elements of antisymmetric matrices are necessarily zero since we must have

$$a_{ii} = -a_{ii} = 0$$

Complex matrices provide interesting and instructive examples of special types. To introduce some of them let us take the complex conjugate of some matrix and then transpose it. For example let

$$A = \begin{bmatrix} a+jb & c+jd \\ d+je & f+jh \end{bmatrix}, \quad \overline{A} = \begin{bmatrix} a-jb & c-jd \\ d-je & f-jh \end{bmatrix}, \quad \overline{A'} = \begin{bmatrix} a-jb & c-je \\ d-jd & f-jh \end{bmatrix},$$

The matrix $\overline{A'}$ is so important in matrix algebra that it is given the special name of *Hermitian adjoint* of A.

Now, it is conceivable that a matrix H exists, with the property that its Hermitian adjoint equals itself. In symbols

$$\overline{H'} = H \tag{2.42}$$

As an example take

$$H = \begin{bmatrix} a & d+jp & e+jq \\ d-jp & b & f+jr \\ e-jq & f-jr & c \end{bmatrix}$$

$$\overline{H} = \begin{bmatrix} a & d-jp & e-jq \\ d+jp & b & f-jr \\ e+jq & f+jr & c \end{bmatrix}$$

$$\overline{H'} = \begin{bmatrix} a & d+jp & e+jq \\ d-jp & b & f+jr \\ e-jq & f-jr & c \end{bmatrix} = H$$

The matrix H is called *Hermitian*. Hermitian matrices are important in many applications, apart from providing practice in transposition, as they do in this section. It is readily seen, on the example above, that a real Hermitian matrix is symmetric.

By analogy with skew symmetric matrices there are *skew Hermitian* matrices. They satisfy the relation

$$\overline{H'} = -H$$

It will be observed that the diagonal elements of a Hermitian matrix are real, since they are equal to their own complex conjugates. The diagonal elements of a skew Hermitian matrix are purely imaginary.

Various types of special matrices occur in problems of partial differentiation. We shall now describe the most important of them, called the *Jacobian matrix*. To introduce it we start with two functions of two independent variables each

$$f_1(x_1, x_2),\ f_2(x_1, x_2)$$

We form partial derivatives of these functions and arrange them in a matrix of order 2×2

$$\begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} \\[2ex] \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} \end{bmatrix} \tag{2.43}$$

This is the Jacobian matrix of the functions f_i with respect to the variables x_j. The method of labelling rows and columns of this matrix should now be noted. The subscripts of the functions in the numerator label the rows, whereas the independent variables label the columns. Abbreviated symbols are frequently used to denote Jacobian matrices. Thus, either $\left[\dfrac{\partial f_i}{\partial x_j}\right]$ or simply $\left[\dfrac{\partial f}{\partial x}\right]$ can stand for (2.43). The latter form is particularly concise and can be used to advantage, as we shall see in a moment.

We shall now obtain a result, expressed in terms of Jacobian matrices, which is analogous to the familiar chain rule of differentiation. We assume that our variables x_j are themselves functions of new independent variables y_k,—$x_j = x_j(y_1, y_2)$, and we wish to find the partial derivatives of the functions f_i with respect to the y_k. Let us first form the product of matrices:

$$\begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} \\[2ex] \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} \end{bmatrix} \begin{bmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2} \\[2ex] \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2} \end{bmatrix}$$

The second matrix of the above product is the Jacobian of the variables x_1, x_2 with respect to the new independent variables y_1, y_2. Written out in full the product is

$$\begin{bmatrix} \dfrac{\partial f_1}{\partial x_1}\dfrac{\partial x_1}{\partial y_1} + \dfrac{\partial f_1}{\partial x_2}\dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial f_1}{\partial x_1}\dfrac{\partial x_1}{\partial y_2} + \dfrac{\partial f_1}{\partial x_2}\dfrac{\partial x_2}{\partial y_2} \\[2ex] \dfrac{\partial f_2}{\partial x_1}\dfrac{\partial x_1}{\partial y_1} + \dfrac{\partial f_2}{\partial x_2}\dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial f_2}{\partial x_1}\dfrac{\partial x_1}{\partial y_2} + \dfrac{\partial f_2}{\partial x_2}\dfrac{\partial x_2}{\partial y_2} \end{bmatrix} \tag{2.44}$$

In terms of typical elements the foregoing manipulations may be written down concisely as follows:

$$\left[\frac{\partial f_i}{\partial x_j}\right]\left[\frac{\partial x_j}{\partial y_k}\right] = \left[\frac{\partial f_i}{\partial x_1}\frac{\partial x_1}{\partial y_k} + \frac{\partial f_i}{\partial x_2}\frac{\partial x_2}{\partial y_k}\right] \tag{2.45}$$

Let us now consider one of the elements of the product matrix (2.44), say the first. It is proved at length in textbooks on the differential calculus that an expression of this form is, in fact, the partial derivative of f_1 with respect to the new variable y_1.

$$\frac{\partial f_1}{\partial x_1}\frac{\partial x_1}{\partial y_1} + \frac{\partial f_1}{\partial x_2}\frac{\partial x_2}{\partial y_1} = \frac{\partial f_1}{\partial y_1} \tag{2.46}$$

A similar observation applies to the remaining elements of (2.44). The product matrix, (2.44) can, therefore, be rewritten

$$\begin{bmatrix} \dfrac{\partial f_1}{\partial y_1} & \dfrac{\partial f_1}{\partial y_2} \\[2ex] \dfrac{\partial f_2}{\partial y_1} & \dfrac{\partial f_2}{\partial y_2} \end{bmatrix} \tag{2.47}$$

The results may be summarised in two alternative forms:

$$\left[\frac{\partial f_i}{\partial x_j}\right]\left[\frac{\partial x_j}{\partial y_k}\right] = \left[\frac{\partial f_i}{\partial y_k}\right]$$

$$\left[\frac{\partial f}{\partial x}\right]\left[\frac{\partial x}{\partial y}\right] = \left[\frac{\partial f}{\partial y}\right] \tag{2.48}$$

The striking analogy between this matrix expression and the chain rule of differentiation is now apparent.

$$\frac{df}{dy} = \frac{df}{dx}\frac{dx}{dy}$$

where $f = f(x)$ and $x = x(y)$.

Equations 2.48 afford perhaps the best example so far of the highly efficient symbolism provided by matrix methods.

2.12 REVERSAL RULE FOR THE TRANSPOSE OF A PRODUCT

Transposition is an algebraic manipulation peculiar to matrices. Its most important consequence appears when an attempt is made to transpose a product of two or more matrices. Is in this connection that the *reversal rule* arises.

Whenever it is necessary to transpose the product of two or more matrices, it is found that the same result is obtained by first transposing the factors, and then multiplying them in reverse order. To put it in symbols we always find

$$C' = (AB)' = B'A' \tag{2.49}$$

The origin of this rule is to be found in the basic definition of matrix multiplication which is row into column, the row being taken from the prefactor and the column from the postfactor. As a result of transposition rows and columns are interchanged and consequently the order of factors must be reversed. Instead of writing out in full a proof of Equation 2.49, we illustrate how the reversal rule works on some numerical examples.

Let

$$A = \begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 4 \\ 2 & -1 \end{bmatrix},$$

Then

$$C = AB = \begin{bmatrix} 6 & 5 \\ 0 & 4 \end{bmatrix} \quad \text{and } C' = (AB)' = \begin{bmatrix} 6 & 0 \\ 5 & 4 \end{bmatrix}$$

$$\text{Now } A' = \begin{bmatrix} 2 & 1 \\ 3 & 0 \end{bmatrix}, \quad B' = \begin{bmatrix} 0 & 2 \\ 4 & -1 \end{bmatrix}$$

$$\text{and } B'A' = \begin{bmatrix} 6 & 0 \\ 5 & 4 \end{bmatrix} = (AB)' = C'$$

The results are seen to agree with Equation 2.49.

Example of a product including a non-square matrix:

$$\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right)' = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix}'$$

$$= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 & a_{21}x_1 + a_{22}x_2 \end{bmatrix}$$

$$= \begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}$$

$$= \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}'\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}'$$

2.13 DETERMINANTS

Determinants enter the subject of matrix algebra when the reciprocal matrix is to be defined. It is assumed that the reader has learned about determinants before, hence the object of this section is to provide a summary of the basic properties for convenient reference, arranged in the context of matrix algebra.

The *determinant* of a *square* matrix A, denoted by the symbol $|A|$, is a scalar number computed from the elements of A. For an array of order 2×2 the method of computation is given by the scheme

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Determinants of higher order are evaluated by expansion in terms of elements of a row or column. A determinant of order 3×3 can be expanded in terms of elements of the first row as follows:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33}\end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23}\\ a_{32} & a_{33}\end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23}\\ a_{31} & a_{33}\end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22}\\ a_{31} & a_{32}\end{vmatrix} \qquad (2.50)$$

Here, the determinants of order 2×2, which multiply the elements of the first row, are called the *minors* of these elements. The minor of an element is obtained by striking out the row and column to which the element belongs. Thus, the minor of a_{12} is

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33}\end{vmatrix} = \begin{vmatrix} a_{21} & a_{23}\\ a_{31} & a_{33}\end{vmatrix}$$

The expansion of a determinant in terms of the elements of the first row and their minors is completed by assigning positive and negative signs to alternate terms, as exemplified in Equation 2.50. The value of the determinant in Equation 2.50 is finally obtained by calculating the minors of order 2×2. The expansion method can be applied to determinants of any order.

It proves more convenient in practice to work in terms of cofactors, or signed minors, when expanding determinants. The *cofactor* of an element a_{ij} is its minor, with a sign prefixed according to the rule $(-1)^{i+j}$. Thus, the cofactor of the element a_{12} considered above is its minor with a negative sign. The cofactor of the element a_{ij} will be denoted by the symbol $|A_{ij}|$ (including sign).

The rule of attaching signs to minors, to form cofactors, is most readily remembered in terms of the following scheme:

$$\begin{bmatrix} + & - & + & - & + & - & + & \cdots\cdots \\ - & + & - & + & - & + & - & \cdots\cdots \\ + & - & + & - & + & - & + & \cdots\cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\cdots \end{bmatrix}$$

The minor of an element is given the sign found in its position in the above array.

In terms of cofactors Equation 2.50 can now be rewritten

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33}\end{vmatrix} =$$

$$= a_{11}(+1)\begin{vmatrix} a_{22} & a_{23}\\ a_{32} & a_{33}\end{vmatrix} + a_{12}(-1)\begin{vmatrix} a_{21} & a_{23}\\ a_{31} & a_{33}\end{vmatrix} + a_{13}(+1)\begin{vmatrix} a_{21} & a_{22}\\ a_{31} & a_{32}\end{vmatrix}$$

or more briefly

$$|A| = a_{11}|A_{11}| + a_{12}|A_{12}| + a_{13}|A_{13}|$$

In general, the determinant $|A|$ may be expanded in terms of the elements of any of its rows, or columns, and their cofactors. The expansion in terms of the elements of the *i-th row* assumes the form

$$|A| = a_{i1}|A_{i1}| + a_{i2}|A_{i2}| + \ldots\ldots + a_{in}|A_{in}|$$
$$= \sum_{l=1}^{n} a_{il}|A_{il}| \qquad (2.51)$$

where $n \times n$, or briefly n, is the order of the matrix A. The expansion in terms of the elements of the j-*th column* is

$$|A| = a_{1j}|A_{1j}| + a_{2j}|A_{2j}| + \ldots\ldots + a_{nj}|A_{nj}|$$
$$= \sum_{k=1}^{n} a_{kj}|A_{kj}| \qquad (2.52)$$

The cofactors appearing in Equations 2.51 and 2.52 may in turn be expanded in terms of their own elements and cofactors of lower order. In fact, the process may be continued until cofactors of order 1×1 are arrived at, and the determinant $|A|$ is expressed as a sum of products of its elements.

Using the method of evaluation outlined above, various properties of determinants may be established. The following is a summary of the properties we shall have occasion to apply.

1. Transposition leaves the value of a determinant unaltered.
2. Interchange of any two rows or columns of a determinant changes its sign.
3. A determinant having two rows or columns identical equals zero.
4. The value of a determinant is multiplied by the factor l, if any one of its rows or columns is multiplied by l.
5. The value of a determinant is unchanged by the addition of multiples of any of its rows (or columns) to a given row (or column).

It should be noted that rule 4 does not apply to the determinant of a matrix multiplied by a scalar k. In that case all rows are multiplied by k, so that

the determinant is multiplied by k^n (n = order of matrix). In symbols

$$|kA| = k^n |A|$$

Using rule 3 we can prove that expansions of a determinant in terms of *alien cofactors* always vanish. By this we mean that substitution into Equation 2.51 of cofactors not belonging to the i-th row makes the expression vanish. Using cofactors of the p-th row, instead of the i-th, we write

$$a_{i1}|A_{p1}| + a_{i2}|A_{p2}| + \dots + a_{in}|A_{pn}| = 0$$

$$\sum_{l=1}^{n} a_{il}|A_{pl}| = 0 \tag{2.53}$$

The truth of Equation 2.53 is evident, once it is realised that the cofactors $|A_{pl}|$ contain the i-th row inside themselves. It is then seen that Equation 2.53 is the expansion of a determinant having two identical rows: the i-th and the p-th.

2.14 THE ADJOINT MATRIX

The cofactors of $|A|$ may be used to form a new matrix from A.

Definition.

Take the cofactors $|A_{ij}|$ and arrange them in the form of a matrix, by replacing each element of A by its cofactor, and then transposing the array. The resulting matrix is called the *adjoint of A*, and is denoted by adj$A$.

In symbols the adjoint of A is defined by the expression

$$\text{adj}A = \left[|A_{ij}|\right]' = \left[|A_{ji}|\right] \tag{2.54}$$

Sometimes the word adjugate is used instead of adjoint.

The adjoint of A should not be confused with the Hermitian adjoint of A; the only thing they have in common is transposition.

In the following example of order 3×3 the successive stages of forming the adjoint matrix should be carefully noted.

Original matrix:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Matrix of cofactors *not* transposed:

$$\left[|A_{ij}|\right] = \begin{bmatrix} |A_{11}| & |A_{12}| & |A_{13}| \\ |A_{21}| & |A_{22}| & |A_{23}| \\ |A_{31}| & |A_{32}| & |A_{33}| \end{bmatrix}$$

$$= \begin{bmatrix} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} & (-1)\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ (-1)\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} & (-1)\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} \\ \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} & (-1)\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \end{bmatrix}$$

Matrix of transposed cofactors or adjoint matrix:

$$\text{adj}A = \left[|A_{ij}|\right]' = \begin{bmatrix} |A_{11}| & |A_{21}| & |A_{31}| \\ |A_{12}| & |A_{22}| & |A_{32}| \\ |A_{13}| & |A_{23}| & |A_{33}| \end{bmatrix}$$

$$= \begin{bmatrix} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} & (-1)\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ (-1)\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} & (-1)\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} \\ \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} & (-1)\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \end{bmatrix}$$

The subscripts of the cofactors in the untransposed and transposed matrices should be carefully noted at this point.

Numerical example:

$$A = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 3 & -1 \\ 4 & 0 & 0 \end{bmatrix}, \quad \left[|A_{ij}|\right] = \begin{bmatrix} 0 & -4 & 12 \\ 0 & 8 & 0 \\ 6 & -1 & 3 \end{bmatrix}, \quad \left[|A_{ij}|\right]' = \text{adj}A = \begin{bmatrix} 0 & 0 & 6 \\ -4 & 8 & -1 \\ 12 & 0 & 3 \end{bmatrix}$$

For a matrix of order 2×2 the adjoint is particularly easy to form since the cofactors are just elements.

Example:

$$Y = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix}, \quad \left[|Y_{ij}|\right] = \begin{bmatrix} y_{22} & -y_{21} \\ -y_{12} & y_{11} \end{bmatrix}, \quad \text{adj}Y = \left[|Y_{ij}|\right]' = \begin{bmatrix} y_{22} & -y_{12} \\ -y_{21} & y_{11} \end{bmatrix}$$

The adjoint matrix is introduced at this stage because it forms an essential step in the definition of the reciprocal matrix.

2.15 THE RECIPROCAL MATRIX

The symbolic similarity between matrix and scalar algebra has already been pointed out. It applies to the reciprocal of a matrix A, which will be found to be analogous to the reciprocal of a scalar number.

This means that the reciprocal matrix, let it be denoted by A^{-1}, should yield the unit matrix when either pre- or postmultiplied by A. Symbolically this is stated as follows:

$$AA^{-1} = A^{-1}A = I \tag{2.55}$$

From the properties of matrix multiplication it follows that Equation 2.55 can only be satisfied by square matrices. Another reason why reciprocals can be formed of square matrices only is the fact that the determinant enters into the definition.

Definition.

To form the *reciprocal* of a square matrix A write down adj$A$ and multiply it by $\frac{1}{|A|}$. In symbols:

$$A^{-1} = \frac{1}{|A|} \text{adj } A \tag{2.56}$$

Example 1.

$$A = \begin{bmatrix} 2 & 0 \\ -3 & 1 \end{bmatrix}, \qquad |A| = 2, \qquad \text{adj}A = \begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix}$$

$$A^{-1} = \frac{1}{|A|}\text{adj}A = 1/2\begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 \\ 3/2 & 1 \end{bmatrix}$$

To check that the last matrix is the reciprocal of A according to the requirements stated above, carry out the multiplications AA^{-1} and $A^{-1}A$.

$$AA^{-1} = \begin{bmatrix} 2 & 0 \\ -3 & 1 \end{bmatrix}\begin{bmatrix} 1/2 & 0 \\ 3/2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

$$A^{-1}A = \begin{bmatrix} 1/2 & 0 \\ 3/2 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

Example 2.

$$A = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 1 & 3 \\ 1 & -2 & 0 \end{bmatrix}, \qquad |A| = 9, \qquad \text{adj}A = \begin{bmatrix} 6 & 0 & -3 \\ 3 & 0 & -6 \\ -1 & 3 & 2 \end{bmatrix}$$

$$A^{-1} = \frac{1}{|A|}\text{adj}A = 1/9\begin{bmatrix} 6 & 0 & -3 \\ 3 & 0 & -6 \\ -1 & 3 & 2 \end{bmatrix}$$

Here the scalar factor $\frac{1}{|A|} = \frac{1}{9}$ is left outside the adjoint as this proves more convenient in numerical computations.

Check:

$$A^{-1}A = 1/9\begin{bmatrix} 6 & 0 & -3 \\ 3 & 0 & -6 \\ -1 & 3 & 2 \end{bmatrix}\begin{bmatrix} 2 & -1 & 0 \\ 0 & 1 & 3 \\ 1 & -2 & 0 \end{bmatrix} = 1/9\begin{bmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{bmatrix} = I$$

$$AA^{-1} = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 1 & 3 \\ 1 & -2 & 0 \end{bmatrix}(1/9)\begin{bmatrix} 6 & 0 & -3 \\ 3 & 0 & -6 \\ -1 & 3 & 2 \end{bmatrix} = 1/9\begin{bmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{bmatrix} = I$$

In the last line the scalar factor has been taken outside the matrix product (see Section 2.7).

In the foregoing examples the reciprocal matrix is seen to satisfy Equation 2.55. It is possible to prove in general that a reciprocal found according to Equation 2.56 always satisfies these requirements. Before writing out a complete proof, however, let us demonstrate this fact for a general matrix of order 3×3.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$A^{-1} = \frac{1}{|A|}\begin{bmatrix} |A_{11}| & |A_{21}| & |A_{31}| \\ |A_{12}| & |A_{22}| & |A_{32}| \\ |A_{13}| & |A_{23}| & |A_{33}| \end{bmatrix}$$

$$A^{-1}A = \frac{1}{|A|}\begin{bmatrix} |A_{11}| & |A_{21}| & |A_{31}| \\ |A_{12}| & |A_{22}| & |A_{32}| \\ |A_{13}| & |A_{23}| & |A_{33}| \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Consider the first element of this product.

$$a_{11}|A_{11}| + a_{21}|A_{21}| + a_{31}|A_{31}| = |A|$$

It is the determinant $|A|$ expanded in terms of the elements of its first column and their cofactors. In a similar way the remaining diagonal elements of the above product can be seen to equal $|A|$. Next let us write out one of the off diagonal elements of the product, say the second element of the first row.

$$a_{12}|A_{11}| + a_{22}|A_{21}| + a_{32}|A_{31}| = 0$$

The sum vanishes since it is an expansion of the determinant $|A|$ in terms of alien cofactors. Similarly all the off diagonal elements can be shown to equal zero. Hence the product has the form

$$A^{-1}A = \frac{1}{|A|}\begin{bmatrix} |A| & 0 & 0 \\ 0 & |A| & 0 \\ 0 & 0 & |A| \end{bmatrix} = I$$

The reversed product AA^{-1} can be shown to equal the unit matrix in the same way.

A general proof that $A^{-1}A = AA^{-1} = I$ follows the lines of the foregoing calculation except that the matrices and their products cannot be written out in full, but must be handled through summations. Let

$$C = AA^{-1} = A\left(\frac{1}{|A|}\text{adj}A\right) \tag{2.57}$$

A typical element of this product is

$$c_{ij} = \frac{1}{|A|}\sum_{k=1}^{n} a_{ik}|A_{jk}| \tag{2.58}$$

where $n \times n$ is the order of A. It should be observed that the summation in Equation 2.58 runs over the column subscript of the cofactors A_{jk} because the adjoint is a transposed matrix. The summation is the expansion of the determinant $|A|$ in terms of elements of its rows. In cases when $j = i$ it yields the value of the determinant, so that we have

$$\sum_{k=1}^{n} a_{ik}|A_{ik}| = |A| \tag{2.59}$$

and the diagonal elements of the product of Equation 2.57 are

$$c_{ii} = 1$$

Whenever $j \neq i$ the summation in Equation 2.58 is an expansion of the determinant $|A|$ in terms of alien cofactors and therefore vanishes. Hence

$$c_{ij} = 0$$

Collecting these results we find that C is a unit matrix.

$$C = I = AA^{-1} \tag{2.60}$$

A similar argument applies to the reversed product $A^{-1}A$. Hence, the reciprocal matrix may be used either in *predivision* or in *postdivision* with A to yield the unit matrix.

A few final observations should be made before passing on to applications of matrix division. Throughout this section it has been tacitly assumed that the determinant $|A|$ does not vanish. Without this assumption it would be impossible to form the scalar factor $\frac{1}{A}$ and hence the reciprocal matrix. There are, of course, square matrices whose determinants vanish and which, therefore, do not have a reciprocal. Such matrices are termed *singular*. Matrices with a nonvanishing determinant are *nonsingular*.

The reciprocal is a matrix like any other, therefore it is subject to the same rules. In particular it does not commute under multiplication by matrices other than A. In general

$$A^{-1}B \neq BA^{-1}$$

For the same reason A^{-1} is never written $\frac{1}{A}$. For example, in the product

$BA^{-1}E$ the sequence of factors would be confused, if the expression were written

$$\frac{BE}{A}$$

2.16 APPLICATIONS OF MATRIX DIVISION

It was mentioned on p. 28 that systems of linear equations can be solved by means of the reciprocal matrix. As a matter of fact, its definition was originally formed with that end in view. It is, therefore, natural that we should solve some linear equations as the first application of matrix division. We consider the 'square' case of n equations in n unknowns.

$$AX = H \tag{2.61}$$

Premultiplication of both sides of Equation 2.61 by A^{-1} yields

$$X = A^{-1}H \tag{2.62}$$

Let us write out in full the case n = 3

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} \tag{2.63}$$

As the first step towards a solution we write

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{|A|}\begin{bmatrix} |A_{11}| & |A_{21}| & |A_{31}| \\ |A_{12}| & |A_{22}| & |A_{32}| \\ |A_{13}| & |A_{23}| & |A_{33}| \end{bmatrix}\begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} \tag{2.64}$$

$$X = \frac{1}{|A|}\,(\mathrm{adj}A)\;H$$

Multiplying out the right hand side of Equation 2.64 we find

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{|A|}\begin{bmatrix} h_1|A_{11}| + h_2|A_{21}| + h_3|A_{31}| \\ h_1|A_{12}| + h_2|A_{22}| + h_3|A_{32}| \\ h_1|A_{13}| + h_2|A_{23}| + h_3|A_{33}| \end{bmatrix} \tag{2.65}$$

By the definition of equality of matrices the individual unknowns are

$$x_1 = \frac{1}{|A|}\left(h_1|A_{11}| + h_2|A_{21}| + h_3|A_{31}|\right)$$

$$x_2 = \frac{1}{|A|}\left(h_1|A_{12}| + h_2|A_{22}| + h_3|A_{32}|\right)$$

$$x_3 = \frac{1}{|A|}\left(h_1|A_{13}| + h_2|A_{23}| + h_3|A_{33}|\right) \tag{2.66}$$

In any specific numerical problem, for which the determinant $|A|$ and the cofactors $|A_{ij}|$ have been computed, Equations 2.66 are the final solution of the system of simultaneous equations (Equation 2.63). However, to obtain a better insight into the nature of the solution, it is desirable to go a step further. An inspection of the expressions in brackets will show that they are expansions of certain related determinants in terms of elements of a column. Thus, the first expression will be recognised as the expansion of the determinant

$$\begin{vmatrix} h_1 & a_{12} & a_{13} \\ h_2 & a_{22} & a_{23} \\ h_3 & a_{32} & a_{33} \end{vmatrix}$$

in terms of the elements of the first column. Likewise the remaining expressions can be identified as the determinants

$$\begin{vmatrix} a_{11} & h_1 & a_{13} \\ a_{21} & h_2 & a_{23} \\ a_{31} & h_3 & a_{33} \end{vmatrix}, \quad \begin{vmatrix} a_{11} & a_{12} & h_1 \\ a_{21} & a_{22} & h_2 \\ a_{31} & a_{23} & h_3 \end{vmatrix}$$

Hence Equations 2.66 can be rewritten in the form

$$x_1 = \frac{\begin{vmatrix} h_1 & a_{12} & a_{13} \\ h_2 & a_{22} & a_{23} \\ h_3 & a_{32} & a_{33} \end{vmatrix}}{|A|}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & h_1 & a_{13} \\ a_{21} & h_2 & a_{23} \\ a_{31} & h_3 & a_{33} \end{vmatrix}}{|A|}, \quad x_3 = \frac{\begin{vmatrix} a_{11} & a_{12} & h_1 \\ a_{21} & a_{22} & h_2 \\ a_{31} & a_{23} & h_3 \end{vmatrix}}{|A|}, \tag{2.67}$$

Equations 2.67 represent Cramer's Rule for the solution of a set of simultaneous linear equations.

With the matrix method three ways of solving sets of linear equations are available: (1) successive elimination of the unknowns; (2) Cramer's Rule; (3) working out the reciprocal matrix and postmultiplying it by the column

vector of constants h_i. Which of these methods should be used will depend on circumstances. In simple cases successive elimination may produce the quickest result. The matrix method should prove advantageous in situations where the adjoint of A happens to be known.

Cases of this kind will be encountered in Chapter 3 in connection with four terminal networks. In problems of a general nature the matrix method is almost always to be preferred for its compactness and lucidity.

In the foregoing problem it was assumed that $|A| \neq 0$, or that the matrix of the equations is non-singular. Systems of equations having a singular matrix A or $|A| = 0$, cannot be solved by the method used above. Neither are we in a position to solve systems of equations having a rectangular matrix (a different number of equations and unknowns, see p. 28). To tackle these more difficult problems additional concepts and methods of matrix algebra are required.

An important problem in simultaneous linear equations arises when the constants on the right hand side of Equation 2.61 are all zero.

$$A\,X = O \tag{2.68}$$

The distinction between equations of the type represented by Equation 2.61 and 2.68 is fundamental, and separate terms are used to describe them. Equation 2.61 is called *inhomogeneous* whereas Equation 2.68 is called *homogeneous*. It is evident that Equation 2.62 does not apply to homogeneous equations, except in the trivial case when all the unknowns are zero. The question of homogeneous equations will be dealt with fully in Chapter 4.

All these limitations need not discourage the reader who wants to learn the applications of matrix algebra to electric circuits. The methods explained in the present chapter will be sufficient to cover many electrical problems, in fact, all those discussed in Chapter 3.

Let us now consider the geometrical interpretation of the reciprocal matrix in linear transformations. From the linear transformation

$$X^{(2)} = AX^{(1)} \tag{2.69}$$

we can form the *reciprocal transformation*

$$X^{(1)} = A^{-1}X^{(2)} \tag{2.70}$$

For point transformations this means that just as A shifts the point $P^{(1)}$ into the point $P^{(2)}$, so A^{-1} returns $P^{(2)}$ to the original position $P^{(1)}$. Graphically this is represented in two dimensions by Fig. 2.7.

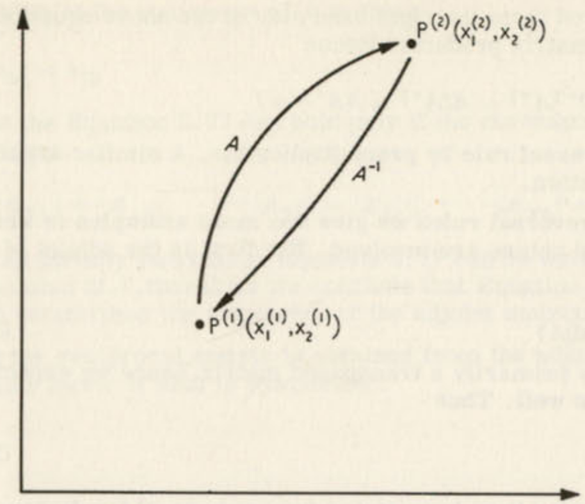


Fig. 2.7

2.17 REVERSAL RULE FOR THE RECIPROCAL OF A PRODUCT MATRIX

A reversal rule applies to the reciprocal of a product of matrices as it does to the transpose of a product (see Section 2.12). This is to be expected since the reciprocal is essentially a transposed matrix. A proof can be given purely in terms of matrix symbols without recourse to summations over elements.

We start by writing down the product of two matrices

$$AB = C \tag{2.71}$$

Since the product C is once again a matrix we can form its reciprocal according to the rules of the preceding section.

$$C^{-1} = (AB)^{-1} \tag{2.72}$$

The question arises whether the reciprocal matrix can be found without first working out Equation 2.71. This is, in fact, possible through the following relation:

$$(AB)^{-1} = B^{-1}A^{-1} \tag{2.73}$$

To verify Equation 2.73 we premultiply it by (AB) and hope to get the unit matrix in the end.

$$CC^{-1} = (AB)(AB)^{-1} = (AB)B^{-1}A^{-1}$$

The brackets are removed from the right hand side of the above equation by the associative law for matrix products. Hence

$$(AB)B^{-1}A^{-1} = ABB^{-1}A^{-1} = AIA^{-1} = AA^{-1} = I$$

This establishes the reversal rule by premultiplication. A similar argument applies to postmultiplication.

While on the subject of reversal rules we give two more examples in which matrices of a transposed nature are involved. The first is the adjoint of a product. We write

$$\mathrm{adj}(AB) = (\mathrm{adj}B)\,(\mathrm{adj}A) \tag{2.74}$$

The Hermitian adjoint is primarily a transposed matrix, hence we expect a reversal rule to apply as well. Thus

$$\overline{(AB)'} = \overline{B'}\,\overline{A'} \tag{2.75}$$

Although the foregoing rules have been stated for products of two factors only, they are easily extended. Thus the reciprocal of ABC is

$$(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$$

Similar extensions apply to the adjoint and Hermitian adjoint of a multiple product.

2.18 SYMMETRIC MATRICES AND THEIR RECIPROCALS

In this section we prove a theorem on the adjoints and reciprocals of symmetric matrices. The theorem is of great importance when matrix algebra is applied to electrical problems, as we shall see in the following chapter.

Theorem

The adjoint and the reciprocal of a symmetric matrix are themselves symmetric matrices.

First we prove the theorem for adjoint matrices. Since the elements of the adjoint matrix, adj$A$, are cofactors of the determinant $|A|$, we must show that they satisfy the relation

$$|A_{ij}| = |A_{ji}| \tag{2.76}$$

To this end we expand $|A|$ in terms of elements of a row and then in terms of elements of the corresponding column. As an example we write down the expansion for the second row and also for the second column.

$$\begin{aligned} |A| &= a_{21}\,|A_{21}| + a_{22}\,|A_{22}| + \cdots + a_{2j}\,|A_{2j}| + \cdots + a_{2n}\,|A_{2n}| = \\ &= a_{12}\,|A_{12}| + a_{22}\,|A_{22}| + \cdots + a_{j2}\,|A_{j2}| + \cdots + a_{n2}\,|A_{n2}| \end{aligned} \tag{2.77}$$

By virtue of the symmetry of A we have

$$a_{2j} = a_{j2}$$

Hence the Equation 2.77 can hold only if the corresponding cofactors are also equal.

$$|A_{21}| = |A_{12}|, \cdots\cdots |A_{2j}| = |A_{j2}|, \cdots\cdots |A_{2n}| = |A_{n2}|$$

Now, an identity like that in Equation 2.77 can be written down for any row and column of A, therefore we conclude that Equation 2.76 holds in general, which establishes the symmetry of the adjoint matrix of A.

Since the reciprocal matrix is obtained from the adjoint on multiplication by a scalar factor it also is symmetric.

Chapter 3

Equations of Linear Two - port Networks

The matrix algebra explained in the preceding chapter was especially developed to handle linear algebraic relations. The equations of linear circuits are themselves linear, hence they provide a natural field for the application of matrix methods. It is the object of the present chapter to show how this is done.

3.1 MESH EQUATIONS OF LUMPED LINEAR CIRCUITS

Applying Kirchoff's Voltage Law to the circuit of Fig. 3.1 the following set of simultaneous linear equations is obtained.

$$\begin{aligned}
&(Z_1 + Z_2 + Z_3)\, I_1 && - Z_3\, I_2 && - Z_2\, I_3 = E_1 \\
&- Z_3\, I_1 + (Z_3 + Z_4 + Z_5)\, I_2 && && - Z_4\, I_3 = E_2 \\
&- Z_2\, I_1 && - Z_4\, I_2 + (Z_2 + Z_4 + Z_6)\, I_3 && = E_3
\end{aligned} \qquad (3.1)$$

The applied e.m.f.s E_1, E_2, E_3 are assumed to be known and the loop currents I_1, I_2, I_3 are the unknowns, to be found by solving the set of simultaneous linear equations 3.1. Our aim is to apply matrix algebra to the solution of this problem.

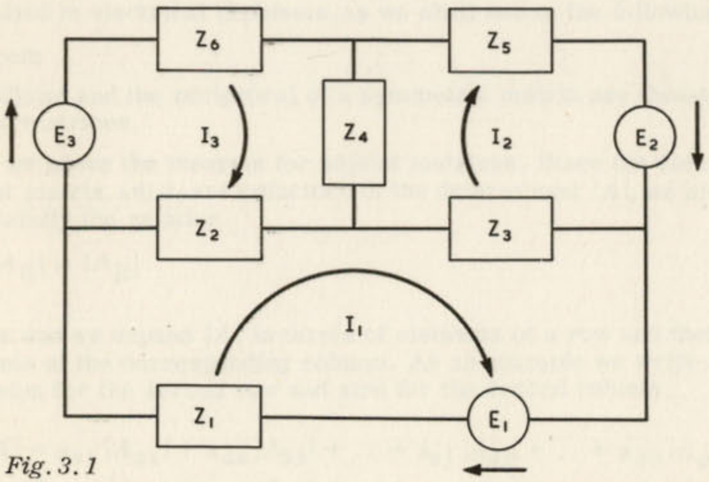


Fig. 3.1

The first step towards a solution is to rewrite Equations 3.1 in a notation better suited to matrices. We replace the separate circuit impedances that make up the matrix of Equations 3.1 by symbols with double subscripts.

$$\begin{bmatrix} Z_1 + Z_2 + Z_3 & -Z_3 & -Z_2 \\ -Z_3 & Z_3 + Z_4 + Z_5 & -Z_4 \\ -Z_2 & -Z_4 & Z_2 + Z_4 + Z_6 \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} & Z_{13} \\ Z_{21} & Z_{22} & Z_{23} \\ Z_{31} & Z_{32} & Z_{33} \end{bmatrix} = Z \qquad (3.2)$$

Equation 3.1 is next rewritten in matrix form.

$$\begin{bmatrix} Z_{11} & Z_{12} & Z_{13} \\ Z_{21} & Z_{22} & Z_{23} \\ Z_{31} & Z_{32} & Z_{33} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} E_1 \\ E_2 \\ E_3 \end{bmatrix}$$

$$Z\, I = E$$

Premultiplication of the last expression by Z^{-1} yields the solution of our problem.

$$I = Z^{-1} E \qquad (3.4)$$

Writing out this result more fully, in accordance with the rule of forming the reciprocal matrix, we find:

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \frac{1}{|Z|} \begin{bmatrix} |Z_{11}| & |Z_{21}| & |Z_{31}| \\ |Z_{12}| & |Z_{22}| & |Z_{32}| \\ |Z_{13}| & |Z_{23}| & |Z_{33}| \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \\ E_3 \end{bmatrix}$$

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} \frac{|Z_{11}|}{|Z|} E_1 + \frac{|Z_{21}|}{|Z|} E_2 + \frac{|Z_{31}|}{|Z|} E_3 \\ \frac{|Z_{12}|}{|Z|} E_1 + \frac{|Z_{22}|}{|Z|} E_2 + \frac{|Z_{32}|}{|Z|} E_3 \\ \frac{|Z_{13}|}{|Z|} E_1 + \frac{|Z_{23}|}{|Z|} E_2 + \frac{|Z_{33}|}{|Z|} E_3 \end{bmatrix} \qquad (3.5)$$

In the above expressions $|Z|$ is the determinant of the matrix defined in Equation 3.2, while $|Z_{ij}|$ is the cofactor of the element Z_{ij}.

Having solved the algebraic problem of finding the loop currents for the three-mesh circuit of Fig. 3.1 let us pause for a while and make a number of observations on the result.

The matrix defined by Equation 3.2 is usually referred to as the *impedance matrix* of the given circuit. Had we chosen to set up the equations of the circuit in Fig. 3.1 by applying Kirchoff's Current Law to the nodes or, in other words, by using the node method, we would have obtained the *admittance matrix* instead.

The elements of the impedance matrix are not simple individual impedances but their sums, sometimes with negative signs. They are nevertheless referred to as impedances, and it is useful to draw a distinction between the diagonal and off diagonal elements by name. The diagonal elements Z_{ii} are called the *self impedances* of the corresponding meshes. They are the sums of the individual circuit impedances taken around a complete loop. The off diagonal elements Z_{ij} are called the *mutual impedances* of loops i and j. They couple electrically the i-th and j-th loops.

It has been tacitly assumed from the outset that the individual circuit impedances are passive, linear, and bilateral or that they are made up of coils, condensers, and linear resistors only. As a result every impedance shared by two loops looks the same, regardless of from which loop it is viewed. Thus the impedance Z_3 is the same to both the current I_1 and I_2. In consequence of this fact the *matrix of Equations 3.1 is symmetric*. In terms of the double subscript symbols introduced by Equation 3.2 this means that

$$Z_{ij} = Z_{ji} \quad \text{or} \quad Z^t = Z \tag{3.6}$$

where Z^t is the transpose of Z.

If some of the inter-mesh impedances included *vacuum valves or transistors the above symmetry relation would not hold.*

An inspection of the solution, Equation 3.5, shows that each mesh current is made up of a sum of terms, each term including one of the e.m.f.s to the first power. This means that each e.m.f. contributes independently towards the current in a given loop. The phenomenon is called the *law of superposition* for passive linear networks, since a loop current is generated by a superposition of the effects of all the e.m.f.s in the circuit.

As a further example of the convenience of the matrix form of circuit equations we prove the *theorem of reciprocity*. This applies to the network of Fig. 3.1 with only one generator present. Assuming at first that $E_2 = E_3 = 0$ and $E_1 = V$ we find the current in mesh 2 from Equation 3.5

$$I_2 = \frac{|Z_{12}|}{|Z|} V \tag{3.7}$$

Next we remove the generator V from loop 1, insert it into loop 2, and calculate the current in loop 1. Putting $E_1 = E_3 = 0$ and $E_2 = V$ into Equation 3.5 we obtain

$$I_1 = \frac{|Z_{21}|}{|Z|} V \tag{3.8}$$

It is proved at length in Section 2.18, that the adjoint, and therefore the reciprocal, of a symmetric matrix is itself symmetric. Hence the coefficients in Equations 3.7 and 3.8 are equal, since they are the appropriate elements of the reciprocal of the symmetric matrix Z. As a result the currents I_2 and I_1, flowing in response to the generator V, placed first in loop 1 and then in loop 2, are equal. This is the theorem of reciprocity, proved for loops 1 and 2. A similar proof can be formulated for each pair of meshes of the circuit, and in every case the argument hinges on the symmetry of the circuit matrix Z and its reciprocal Z^{-1}.

Since the reciprocity theorem depends on the symmetry of the impedance matrix Z, it does *not apply to circuits containing valves or transistors*.

So far matrix methods have been applied to a comparatively simple circuit consisting of three meshes only. The main advantage of matrix algebra lies in the fact that it makes possible a clear and simple formulation of perfectly general problems. To illustrate this advantage we now write down the equations of a general network of n meshes. Direct application of Kirchoff's Laws yields n equations having the form of Equations 3.1. The sums of individual impedances that make up the coefficients of this system of linear equations are replaced by symbols with double subscripts as in Equation 3.2. As a result the circuit equations can be written in the following form:

$$\begin{aligned}
&Z_{11} I_1 + Z_{12} I_2 + \cdots\cdots + Z_{1n} I_n = E_1 \\
&Z_{21} I_1 + Z_{22} I_2 + \cdots\cdots + Z_{2n} I_n = E_2 \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&Z_{n1} I_1 + Z_{n2} I_2 + \cdots\cdots + Z_{nn} I_n = E_n
\end{aligned} \tag{3.9}$$

In matrix form these are

$$\begin{bmatrix} Z_{11} & Z_{12} \cdots\cdots & Z_{1n} \\ Z_{21} & Z_{22} \cdots\cdots & Z_{2n} \\ \cdots & \cdots\cdots\cdots & \cdots \\ Z_{n1} & Z_{n2} & Z_{nn} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_n \end{bmatrix} = \begin{bmatrix} E_1 \\ E_2 \\ \vdots \\ E_n \end{bmatrix} \tag{3.10}$$

$$ZI = E$$

The constants E_i on the right hand side are vector sums of e.m.f.s taken round respective loops, the unknowns I_j are loop currents, the diagonal elements of the matrix Z are self impedances of the loops, and the off diagonal elements are mutual impedances between the loops. The impedance matrix

Z of the circuit is symmetric for networks containing passive resistors, coils, and condensers only. Whenever valves or transistors are included the matrix Z is not symmetric.

It should be noted that in this chapter the symbol I is always used to denote a column vector of currents and not the unit matrix. The latter will be denoted by the symbol U.

The solution of Equation 3.10 is

$$I = Z^{-1}E \tag{3.11}$$

Written out fully in the form analogous to Equations 3.5 this is

$$\begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_n \end{bmatrix} = \frac{1}{|Z|}\begin{bmatrix} |Z_{11}|E_1 + |Z_{21}|E_2 + \cdots\cdots + |Z_{n1}|E_n \\ |Z_{12}|E_1 + |Z_{22}|E_2 + \cdots\cdots + |Z_{n2}|E_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ |Z_{1n}|E_1 + |Z_{2n}|E_2 + \cdots\cdots + |Z_{nn}|E_n \end{bmatrix} \tag{3.12}$$

Equation 3.12 can be made the basis for proving circuit theorems by an extension of the methods used above to demonstrate superposition and reciprocity for the network of Fig. 3.1.

3.2 THE ADMITTANCE PARAMETERS OF TWO-PORT NETWORKS

In many situations it is unnecessary to know the internal structure of a circuit in detail. It may be sufficient to replace a complicated network by a 'black box' leaving two pairs of terminals accessible for the insertion of generators, loads, or measuring instruments. Such a simplified circuit is then termed a *four terminal network* or *two-port network*.

To begin with let us consider the passive circuit of Fig. 3.1 with the generator E_3 removed. Loops 1 and 2 are singled out for attention as the two accessible ports to which generators are connected. For convenience the circuit is redrawn as shown in Fig. 3.2.

The generators in loops 1 and 2 have been renamed V_1 and V_2 to conform with the notation used in connection with two-port networks. As we are interested in the terminal loops only, we pick out of Equation 3.5 of the preceding section the expressions for the currents I_1 and I_2.

$$I_1 = \frac{|Z_{11}|}{|Z|}V_1 + \frac{|Z_{21}|}{|Z|}V_2$$

$$I_2 = \frac{|Z_{12}|}{|Z|}V_1 + \frac{|Z_{22}|}{|Z|}V_2 \tag{3.13}$$

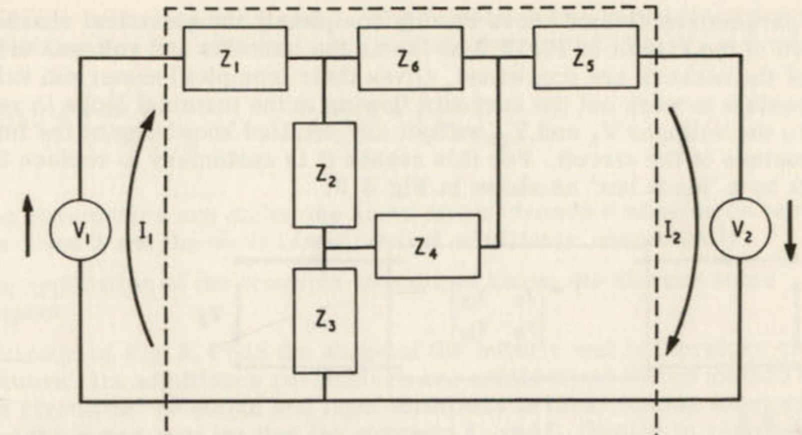


Fig. 3.2

The coefficients in this pair of equations can be obtained from the impedance matrix of the circuit. However, in practical cases this would be cumbersome, and it will be explained below that they can be found by a simple measurement. As they have the dimensions of admittance, it is customary to replace them by single letter symbols with double subscripts, as follows:

$$\frac{|Z_{11}|}{|Z|} = y_{11}, \quad \frac{|Z_{21}|}{|Z|} = y_{12}$$

$$\frac{|Z_{12}|}{|Z|} = y_{21}, \quad \frac{|Z_{22}|}{|Z|} = y_{22} \tag{3.14}$$

The coefficients defined by Equations 3.14 are called the *admittance parameters* or y-*parameters* of a two-port network. It should be noted that the double subscripts do not all agree. This is so because the $|Z_{12}|$ and $|Z_{21}|$ are transposed cofactors used in the formation of the reciprocal matrix Z^{-1}, whereas the symbols y_{12} and y_{21} are defined to form a new matrix. Equations 3.13 can now be rewritten in matrix form using the new symbols.

$$I = YV$$

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix}\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \tag{3.15}$$

In full

$$I_1 = y_{11}V_1 + y_{12}V_2$$
$$I_2 = y_{21}V_1 + y_{22}V_2 \tag{3.16}$$

The y-parameters defined above specify completely the electrical characteristics of the circuit of Fig. 3.2 as far as the currents and voltages in the ports of the network are concerned. Given their (complex) numerical values it is possible to work out the currents flowing in the terminal loops in response to the voltages V_1 and V_2, without any detailed knowledge of the internal structure of the circuit. For this reason it is customary to replace the network by a 'black box' as shown in Fig. 3.3.

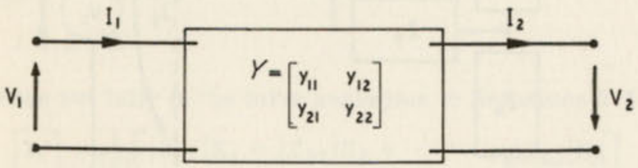


Fig. 3.3

Although the foregoing results have been obtained for the simple circuit of Figs. 3.1 and 3.2, they apply to the general network of n meshes discussed at the end of the preceding section. Using an analogous argument we apply e.m.f.s V_1 and V_2 to the general circuit and extract the expressions for the currents in loops 1 and 2 from Equation 3.12. These are of the same form as Equation 3.13. Hence any circuit, considered as a two port network, is fully specified by four admittance parameters, and can be replaced by the 'black box' of Fig. 3.3.

Let us next consider the procedure whereby the y-parameters can be measured or calculated directly for simple circuits. At first loop 2, or the output loop, is short circuited so that $V_2 = 0$. In this case Equations 3.16 reduce to the simple forms

$$I_1 = y_{11} V_1$$
$$I_2 = y_{21} V_1 \qquad (3.17)$$

Alternatively we can write

$$y_{11} = \left(\frac{I_1}{V_1}\right)_{V_2=0}$$
$$y_{21} = \left(\frac{I_2}{V_1}\right)_{V_2=0} \qquad (3.18)$$

Hence y_{11} is the ratio of the current and voltage in the input loop, or loop 1, while the output port is short circuited. Similarly y_{21} is the ratio of the current in the output port to the voltage in the input port, while the output is short circuited. These facts suggest that y_{11} should be termed the *short circuit input admittance* of the 'black box' and y_{21} the *short circuit transfer admittance* between ports 2 and 1. The remaining two parameters are found in a similar way, but with the input port shortcircuited. For $V_1 = 0$ we find

$$y_{12} = \left(\frac{I_1}{V_2}\right)_{V_1=0} \qquad y_{22} = \left(\frac{I_2}{V_2}\right)_{V_1=0} \qquad (3.19)$$

These parameters are called the *short circuit transfer admittance* between ports 1 and 2 and the *short circuit output admittance* respectively.

As an application of the concepts introduced above we discuss some examples.

The circuit of Fig. 3.4 has the shape of the letter π and is therefore called a π-network. Its admittance parameters are easily found by the method of short circuiting the output and input terminals in turn. Having short circuited the output port we find the currents I_1 and I_2 flowing in response to the input voltage V_1.

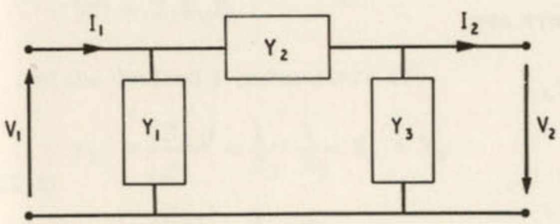


Fig. 3.4

$$Y = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} = \begin{bmatrix} Y_1 + Y_2 & Y_2 \\ Y_2 & Y_2 + Y_3 \end{bmatrix}$$

The circuit now appears as in Fig. 3.5. It is seen that

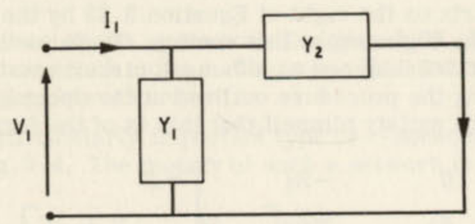


Fig. 3.5

$$I_1 = (Y_1 + Y_2) V_1 = y_{11} V_1$$

and $y_{11} = Y_1 + Y_2$ $\qquad (3.20)$

Also $I_2 = Y_2V_1 = y_{21}V_1$

and $y_{21} = Y_2$ (3.21)

Next the input terminals of the π-network are short circuited and the currents due to the voltage V_2 in loop 2 are calculated with the help of Fig. 3.6.

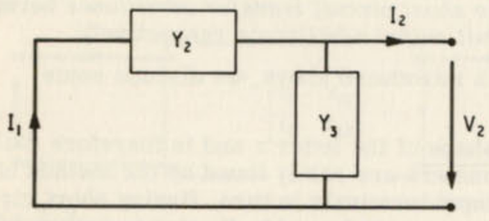


Fig. 3.6

The remaining two y-parameters are

$$\left(\frac{I_2}{V_2}\right)_{V_1=0} = y_{22} = Y_2 + Y_3$$

$$\left(\frac{I_1}{V_2}\right)_{V_1=0} = y_{12} = y_2 \qquad (3.22)$$

Collecting the above results we can write down in full the matrix of y-parameters for the network of Fig. 3.4.

$$Y = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} = \begin{bmatrix} Y_1 + Y_2 & Y_2 \\ Y_2 & Y_2 + Y_3 \end{bmatrix} \qquad (3.23)$$

It is instructive to obtain the matrix on the right of Equation 3.23 by the alternative method explained at the beginning of this section. To do so the impedance matrix of the π-network considered as a 3-mesh circuit must first be written down. By following the procedure outlined at the opening of the present chapter the reader will satisfy himself that this is of the form

$$\begin{bmatrix} Z_{11} & Z_{12} & Z_{13} \\ Z_{21} & Z_{22} & Z_{23} \\ Z_{31} & Z_{32} & Z_{33} \end{bmatrix} = \begin{bmatrix} Z_1 & 0 & -Z_1 \\ 0 & Z_3 & -Z_3 \\ -Z_1 & -Z_3 & Z_1 + Z_2 + Z_3 \end{bmatrix} = Z$$

For convenience impedances have been written for the circuit elements of Fig. 3.4. To find the y-parameters from Equations 3.14, the determinant

and some of the cofactors of the above matrix are required. Expanding the determinant $|Z|$ in terms of the elements of the first row and their cofactors we find

$$|Z| = Z_1|Z_{11}| - Z_1|Z_{13}| = Z_1\left(|Z_{11}| - |Z_{13}|\right)$$

The following five cofactors are needed:

$$|Z_{11}| = Z_3\left(Z_1 + Z_2\right)$$
$$|Z_{12}| = Z_1 Z_3$$
$$|Z_{21}| = Z_1 Z_3 = |Z_{12}|$$
$$|Z_{22}| = Z_1\left(Z_2 + Z_3\right)$$
$$|Z_{13}| = Z_1 Z_3$$

Hence

$$|Z| = Z_1Z_2Z_3$$

and the desired y-parameters are

$$y_{11} = \frac{|Z_{11}|}{|Z|} = \frac{1}{Z_1} + \frac{1}{Z_2} = Y_1 + Y_2$$

$$y_{12} = \frac{|Z_{21}|}{|Z|} = \frac{1}{Z_2} = Y_2$$

$$y_{21} = \frac{|Z_{12}|}{|Z|} = y_{12} = Y_2$$

$$y_{22} = \frac{|Z_{22}|}{|Z|} = \frac{1}{Z_2} + \frac{1}{Z_3} = Y_2 + Y_3$$

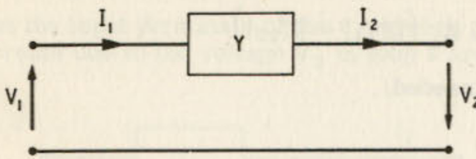
These results agree with Equation 3.23. It is clear that the above calculation is much more cumbersome than the method of short circuiting the ports of the network in turn.

A particularly important type of π-network is obtained when $Y_3 = Y_1$ in Fig. 3.4. The matrix of such a network is

$$\begin{bmatrix} Y_1 + Y_2 & Y_2 \\ Y_2 & Y_1 + Y_2 \end{bmatrix} \qquad (3.24)$$

Our next example is extremely simple. It consists of a single admittance connected in series between the input and output ports as shown in Fig. 3.7.

Its y-parameters are easily written down using the short circuiting technique and noting that $I_2 = I_1$ always.



$$Y = \begin{bmatrix} Y & Y \\ Y & Y \end{bmatrix}$$

Fig. 3.7

$$\begin{aligned}
\left(\frac{I_1}{V_1}\right)_{V_2=0} &= Y = y_{11} \\
\left(\frac{I_2}{V_1}\right)_{V_2=0} &= \left(\frac{I_1}{V_1}\right)_{V_2=0} = Y = y_{21} \\
\left(\frac{I_2}{V_2}\right)_{V_1=0} &= Y = y_{22} \\
\left(\frac{I_1}{V_2}\right)_{V_1=0} &= \left(\frac{I_2}{V_2}\right)_{V_1=0} = Y = y_{12}
\end{aligned} \tag{3.25}$$

One equation is sufficient to describe this circuit:

$$I_1 = Y\left(V_1 + V_2\right)$$

It can be written down immediately by applying Kirchoff's Laws. Nevertheless, we shall find it advantageous in some problems to use the full matrix found above.

It may be instructive to note that the matrix of the network of Fig. 3.7 can be obtained as a limiting case of the π-network. Putting $Y_1 = Y_3 = 0$ in Equation 3.23 and redefining $Y_2 = Y$, Equation 3.25 results.

By reference to Equation 3.23 the Y matrix of the π-network can be seen to be symmetric. This observation applies to the admittance parameters of all two-port networks made up of passive bilateral circuit elements. The general fact follows from Equations 3.14, since the y-parameters are equal to the elements of the reciprocal matrix Z^{-1}, which was shown to be symmetric in Section 3.1. The Y matrix is not symmetric for two port networks that include active or unilateral elements such as valves or transistors.

3.3 THE IMPEDANCE PARAMETERS OF TWO-PORT NETWORKS

In the foregoing section the currents in the terminal loops of a two-port network have been expressed in terms of the terminal voltages with the help of the admittance parameters. This development followed naturally from the formulation of the mesh equations of a circuit in Section 3.1. In many practical problems the position is reversed in that it is necessary to have the voltages expressed as linear functions of the currents. Relations applicable in such situations are obtained by treating V_1 and V_2 as unknowns in Equation 3.15 and solving them.

$$V = Y^{-1}I$$

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \frac{1}{|Y|}\begin{bmatrix} y_{22} & -y_{12} \\ -y_{21} & y_{11} \end{bmatrix}\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \tag{3.26}$$

In plain algebraic form the above equations are

$$V_1 = \frac{y_{22}}{|Y|}\,I_1 - \frac{y_{12}}{|Y|}\,I_2$$

$$V_2 = -\frac{y_{21}}{|Y|}\,I_1 + \frac{y_{11}}{|Y|}\,I_2$$

The coefficients have the dimensions of impedance and are therefore called the *impedance parameters or z-paramèters* of a two-port network. They are usually denoted by lower index letters with double subscripts, by analogy with the y-parameters.

$$\frac{y_{22}}{|Y|} = z_{11},\ -\frac{y_{12}}{|Y|} = z_{12}$$

$$-\frac{y_{21}}{|Y|} = z_{21},\quad \frac{y_{11}}{|Y|} = z_{22} \tag{3.27}$$

The subscripts of the y-parameters and z-parameters do not agree because the latter are defined to form a new matrix. Equation 3.26 can now be rewritten in terms of the new matrix as follows:

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix}\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \tag{3.28}$$

$$V = ZI$$

The symbol Z used here should be carefully distinguished from the same symbol used in Section 3.1. There it stood for the self and mutual

impedances of the individual circuit loops, whereas here it represents the matrix of impedance parameters of a two port network. The latter are related to the individual circuit impedances in a complicated way, which can be traced back through Equations 3.14 and 3.27.

The parameters z_{11} and z_{22} are called the *open circuit input* and *output impedances* of a two-port network, respectively. The parameters z_{12} and z_{21} are the *open circuit transfer impendances*. These names are analogous to those used in connection with the y-parameters in the preceding section. To obtain the z-parameters the ports of a network are open circuited in turn. Starting with the output port we have $I_2 = 0$, whence Equation 3.28 simplifies to the expressions

$$z_{11} = \left(\frac{V_1}{I_1}\right)_{I_2=0}$$

$$z_{21} = \left(\frac{V_2}{I_1}\right)_{I_2=0} \tag{3.29}$$

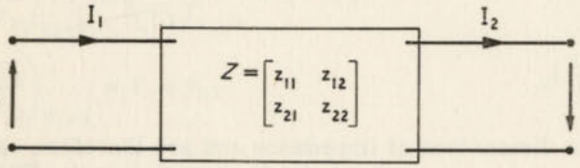


Fig. 3.8

Open circuiting the input port $I_1 = 0$ and we are led to the relations

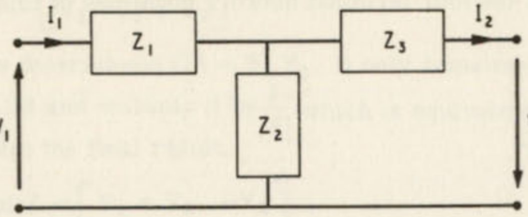
$$z_{12} = \left(\frac{V_1}{I_2}\right)_{I_1=0}$$

$$z_{22} = \left(\frac{V_2}{I_2}\right)_{I_1=0} \tag{3.30}$$

Let us now calculate the z-parameters of some specific circuits. Arranging three impedances in the form of the letter T we obtain the T-network of Fig. 3.9. Leaving the output terminals open circuited we find the relations between the current in loop 1 and the input and output voltages as defined by Equations 3.29 above.

$$z_{11} = \left(\frac{V_1}{I_1}\right)_{I_2=0} = Z_1 + Z_2$$

$$z_{21} = \left(\frac{V_2}{I_1}\right)_{I_2=0} = -Z_2$$



$$Z = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} = \begin{bmatrix} Z_1 + Z_2 & -Z_2 \\ -Z_2 & Z_3 + Z_2 \end{bmatrix}$$

Fig. 3.9

The results are summarised by the matrix in Fig. 3.9.

The negative sign takes account of the *assumed* polarity of V_2 which is opposite to the voltage across the impedance Z_2, when the current I_1 flows through it in the direction shown. In symbols $V_2 = -Z_2 I_1$. Performing the same calculation with the input open circuited we find.

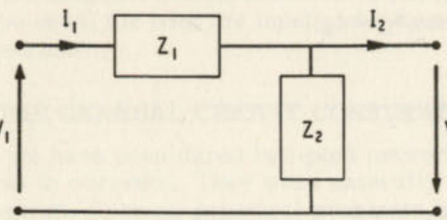
$$z_{12} = \left(\frac{V_1}{I_2}\right)_{I_1=0} = -Z_2 \qquad z_{22} = \left(\frac{V_2}{I_2}\right)_{I_1=0} = Z_3 + Z_2$$

When $Z_3 = Z_1$ the matrix of the T-network simplifies to the form

$$\begin{bmatrix} Z_1 + Z_2 & -Z_2 \\ -Z_2 & Z_1 + Z_2 \end{bmatrix} \tag{3.31}$$

Two impedances arranged in the form of an inverted letter L provide our next example. The z-parameters of the L-network are easily calculated using the open circuiting procedure. Referring to Fig. 3.10 we find

$$z_{11} = Z_1 + Z_2, \quad z_{12} = -Z_2$$
$$z_{21} = \; -Z_2, \quad z_{22} = \; Z_2$$



$$Z = \begin{bmatrix} Z_1 + Z_2 & -Z_2 \\ -Z_2 & Z_2 \end{bmatrix}$$

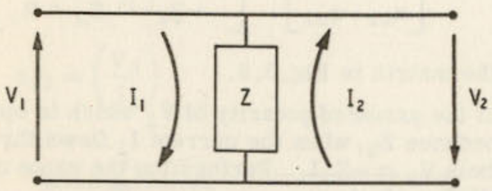
Fig. 3.10

Finally we consider a network consisting of a single impedance, coupling the input and output ports. Working through the open circuiting procedure the reader will satisfy himself that its impedance parameters are as shown in Fig. 3.11. Written out in full the four terminal network equations of this circuit are

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} Z & -Z \\ -Z & Z \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \tag{3.32}$$

$$V_1 = Z(I_1 - I_2)$$

$$V_2 = Z(I_2 - I_1)$$



$$Z = \begin{bmatrix} Z & -Z \\ -Z & Z \end{bmatrix}$$

Fig. 3.11

The last form can be obtained immediately on applying Kirchoff's Current Law to the circuit of Fig. 3.11.

In some problems the y-parameters of a two-port network are already known, but the z-parameters are required. In such cases Equations 3.27 can be used to evaluate the impedance matrix. To illustrate this possibility let us check that the Z matrix of the L-network is as given in Fig. 3.10, by first finding its admittance parameters and then using the relation $Z = Y^{-1}$. Applying the short circuiting procedure we obtain

$$y_{11} = \left(\frac{I_1}{V_1}\right)_{V_2=0} = \frac{1}{Z_1} = Y_1$$

$$y_{21} = \left(\frac{I_2}{V_1}\right)_{V_2=0} = \frac{1}{Z_2} = Y_1 \text{ (since } I_2 = I_1)$$

$$y_{12} = \left(\frac{I_1}{V_2}\right)_{V_1=0} = \left(\frac{I_1}{I_1 Z_1}\right)_{V_1=0} = Y_1$$

$$y_{22} = \left(\frac{I_2}{V_2}\right)_{V_1=0} = \frac{1}{Z_1} + \frac{1}{Z_2} = Y_1 + Y_2$$

Hence the Y matrix of the L-network is

$$Y = \begin{bmatrix} Y_1 & Y_1 \\ Y_1 & Y_1 + Y_2 \end{bmatrix} \tag{3.33}$$

and its determinant $|Y| = Y_1 Y_2$. It only remains to form the adjoint of Equation 3.33 and multiply it by $\frac{1}{|Y|}$, which is equivalent to using Equations 3.27, to obtain the final result.

$$\text{adj}Y = \begin{bmatrix} Y_1 + Y_2 & -Y_1 \\ -Y_1 & Y_1 \end{bmatrix}$$

$$Y^{-1} = \frac{1}{|Y|}\text{adj}Y = \begin{bmatrix} Z_1 + Z_2 & -Z_2 \\ -Z_2 & Z_2 \end{bmatrix} = Z$$

The result agrees with Fig. 3.10.

It will be realised that the foregoing method can be applied in reverse. Given the impedance parameters of a network the Y matrix can be obtained by working out the reciprocal Z^{-1}, instead of using the short circuiting procedure.

In all the examples the matrices of z-parameters can be seen to be symmetric. The observation applies to the Z matrices of all two-port networks composed of passive circuit elements. To prove this statement in general we refer back to Equations 3.27, which define the impedance parameters in terms of the y-parameters. Since it was shown in the preceding section that the latter form a symmetric matrix, it follows that the z_{ij} are also symmetric. The matrices of networks containing valves or transistors or other unilateral elements, such as ferrite isolators, are not symmetric.

The matrices of some of the circuits discussed above display an additional regularity in their structure. The diagonal elements of the matrices of Figs. 3.7 and 3.11 are equal. The same applies to the matrices in Equation 3.24 and 3.31 of the *symmetric* π- and T-networks. The networks are said to be symmetric because they look the same whether viewed from the input or output end. The last remark provides an explanation of the phenomenon. Since the diagonal elements are terminal impedances and admittances, they must be equal for both the input and output ports of all networks which are symmetric.

3.4 THE GENERAL CIRCUIT CONSTANTS

So far we have considered two-port network parameters relating terminal voltages to currents. They were naturally called impedance and admittance parameters. In many practical problems it is necessary to group the terminal electrical quantities V_1, V_2, I_1, I_2 in ways different from those discussed

in the preceding sections. One of the most important arrangements is to have the input quantities V_1, I_1 on one side of a pair of equations and the output quantities V_2, I_2 on the other. The required relations are obtained by solving the second of Equations 3.16 for V_1, and then substituting into the first.

$$V_1 = -\frac{y_{22}}{y_{21}}V_2 + \frac{1}{y_{21}}I_2$$

$$I_1 = -\left(-y_{12} + \frac{y_{11}\,y_{22}}{y_{21}}\right)V_2 + \frac{y_{11}}{y_{21}}I_2 \tag{3.34}$$

The complicated coefficients in this pair of equations are next replaced by single letter symbols, and the equations themselves are rewritten in the following form:

$$V_1 = -AV_2 + BI_2 = AV_2' + BI_2$$

$$I_1 = -CV_2 + DI_2 = CV_2' + DI_2$$

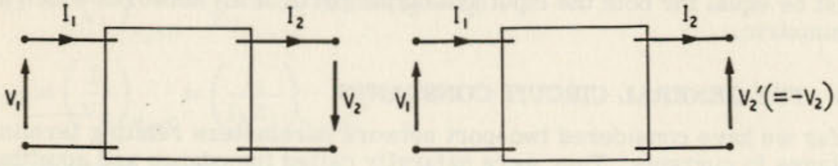
$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} -V_2 \\ I_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} V_2' \\ I_2 \end{bmatrix} \tag{3.35}$$

The negative sign used with V_2 implies a change in the assumed polarity of the output voltage, which is explained by Fig. 3.12. The change of sign convention is made for reasons which will become apparent in the section dealing with cascade connections of four terminal networks.

The coefficients A, B, C, D are called the *general circuit constants*. They express the voltage and current at the input of a four terminal network as a linear transformation of the voltage and current at the output. We agree to write Equations 3.35, using single letter matrix symbols, in the form

$$W_1 = AW_2 \tag{3.36}$$

W_1 and W_2 are column vectors of input and output electrical quantities, and A will be referred to briefly as the A— matrix of a two-port network. Although



Original sign convention *Fig. 3.12* *Modified sign convention used with A, B, C, D*

we shall use the older notation A, B, C, D for the general circuit constants, it should be noted that the proper matrix notation

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

is used more and more often at present.

The general circuit constants can be expressed in terms of the y-parameters using Equations 3.34 and 3.35. The latter are in turn defined in terms of the basic impedance elements of a circuit by Equations 3.14 Although the A— matrix can be found by carrying out this two-stage calculation, in most cases the procedure would prove cumbersome. It is preferable to apply again the method of open circuiting or short circuiting the ports, which proved so convenient in the evaluation of the y- and z- parameters. Open circuiting the output $I_2 = 0$, and the constants A and C are easily found; short circuiting it $V_2' = 0$, and B and D are obtained.

$$A = \left(\frac{V_1}{V_2'}\right)_{I_2=0}, \qquad B = \left(\frac{V_1}{I_2}\right)_{V_2'=0},$$

$$C = \left(\frac{I_1}{V_2'}\right)_{I_2=0}, \qquad D = \left(\frac{I_1}{I_2}\right)_{V_2'=0}, \tag{3.37}$$

It should be noted that in the present case we look into the input port all the time while alternately short circuiting and open circuiting the output terminals. The general circuit constants are a mixture of short circuit and open circuit parameters, in contrast to the short circuit y-parameters and open circuit z-parameters. The physical significance of the A matrix is therefore not so clear, although it is worth noting that the constant A is an open circuit voltage ratio. It can be identified with the amplification factor of a vacuum valve and the voltage transformation ratio of an ideal transformer.

Some examples of circuits and their A matrices will next be given. First the series impedance of Fig. 3.7. Applying Equations 3.37 and remembering that $V_2' = -V_2$ it is found that

$A = 1$ since $V_2' = V_1$,

$B = Z$ since $I_2 = I_1$,

$C = 0$ since $I_1 = I_2 = 0$,

$D = 1$ since $I_2 = I_1$.

Hence the equations relating the input and output of this network are

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} 1 & Z \\ 0 & 1 \end{bmatrix}\begin{bmatrix} V_2' \\ I_2 \end{bmatrix} \tag{3.38}$$

Next let us write down the A-matrix of the shunt element of Fig. 3. 11.

$A = 1$ since $V_2' = V_1$,

$B = 0$ since on short circuit a finite voltage V_1 will cause an infinite current $I_1 = I_2$,

$C = Y$ since $V_2' = V_1$,

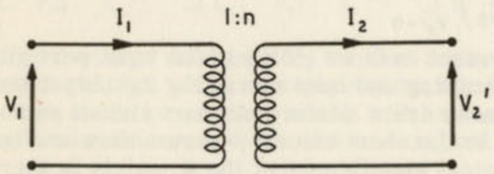
$D = 1$ since $I_2 = I_1$ on short circuit.

Hence the equations are

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ Y & 1 \end{bmatrix} \begin{bmatrix} V_2' \\ I_1 \end{bmatrix} \tag{3. 39}$$

An instructive example of the general circuit constants of a two-port network is provided by the ideal transformer of Fig. 3. 13. On open circuit the terminal voltage ratio equals the turns ratio,

$$A = \left(\frac{V_1}{V_2'}\right)_{I_2=0} = \frac{1}{n}$$



$$A = \begin{bmatrix} \frac{1}{n} & 0 \\ 0 & n \end{bmatrix}$$

Fig. 3. 13

On short circuit the current ratio equals the reciprocal of the turns ratio,

$$D = \left(\frac{I_1}{I_2}\right)_{V_2'=0} = n$$

Still on short circuit, there will be an infinite current in the secondary, or output port, of the transformer in response to a finite voltage applied to the primary. Hence

$$B = 0$$

With the secondary open circuited no current flows in the primary winding of an ideal transformer, hence

$$C = 0$$

The results are collected in the matrix of Fig. 3. 13, and the equations connecting the input and output of an ideal transformer are

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} \frac{1}{n} & 0 \\ 0 & n \end{bmatrix} \begin{bmatrix} V_2' \\ I_2 \end{bmatrix} \tag{3. 40}$$

$$V_1 = \frac{1}{n} V_2'$$

$$I_1 = n I_2$$

The transformer applies a linear transformation between its terminal voltages and currents, as represented by Equations 3. 40.

Our final example of general circuit constants is taken from transmission line theory. We recall that the voltages and currents at two points of a transmission line are related by the equations

$$\begin{aligned} V_1 &= V_2' \cosh \gamma l + I_2 Z_0 \sinh \gamma l \\ I_1 &= V_2' Y_0 \sinh \gamma l + I_2 \cosh \gamma l \end{aligned} \tag{3. 41}$$

where

l = length of line under consideration;

γ = propagation constant of line;

$Z_0 = \frac{1}{Y_0}$ = characteristic impedance of line

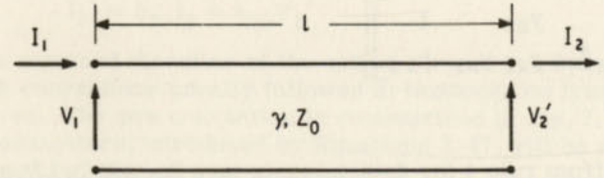


Fig. 3. 14

The length of line shown in Fig. 3. 14 is a type of two-port network, its equations expressed in terms of the general circuit constants being given by Equations 3. 41. Rewriting Equations 3. 41 in matrix form we find

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} \cosh \gamma l & Z_0 \sinh \gamma l \\ Y_0 \sinh \gamma l & \cosh \gamma l \end{bmatrix} \begin{bmatrix} V_2' \\ I_2 \end{bmatrix} \tag{3. 42}$$

In many problems it is more convenient to characterise a section of line by its electrical length, $\theta = \gamma l$, rather than by the propagation constant and physical length separately. In such cases the A matrix is

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \cosh\theta & Z_0 \sinh\theta \\ Y_0 \sinh\theta & \cosh\theta \end{bmatrix} \qquad (3.43)$$

It should be noted that the change of sign convention regarding the output voltage, which was made at the opening of this section, is appropriate to transmission line problems. Equations 3.41 are always written using this convention.

So far we have introduced the equations expressing the input electrical quantities of a two-port network as linear functions of the output quantities. There are occasions when it is desirable to reverse the position by solving Equations 3.35 for the latter. In matrix form the solution is

$$\begin{bmatrix} V_2' \\ I_2 \end{bmatrix} = \frac{1}{|A|} \begin{bmatrix} D & -B \\ -C & A \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix} \qquad (3.44)$$

For passive networks, consisting exclusively of bilateral circuit elements, Equations 3.44 assume a particularly simple form since $|A| = 1$. To prove this we refer back to Equations 3.34, which define the A matrix in terms of the y-parameters, and recall from Section 3.2 that the latter are symmetric. Hence

$$|A| = \left| \frac{1}{y_{21}} \begin{bmatrix} y_{22} & 1 \\ -y_{21}^2 + y_{11} y_{22} & y_{11} \end{bmatrix} \right|$$

$$= \left(\frac{1}{y_{21}} \right)^2 \left| \begin{bmatrix} y_{22} & 1 \\ -y_{21}^2 + y_{11} y_{22} & y_{11} \end{bmatrix} \right|$$

$$= 1$$

The second line follows from rule 4 for determinants (see Section 2.13), since both columns or rows are multiplied by the factor $1/y_{12}$. Hence for bilateral networks

$$A = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \text{ and } A^{-1} = \begin{bmatrix} D & -B \\ -C & A \end{bmatrix} \qquad (3.45)$$

The A matrix of the transmission line, Equation 3.43, affords a particularly good example of this relation, but the reader will, no doubt, check it for the remaining circuits discussed above.

The reciprocal matrix A^{-1} is frequently denoted by

$$A^{-1} = B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

3.5 THE h- AND g- PARAMETERS OF TWO-PORT NETWORKS

The terminal electrical quantities of a two-port network have been grouped in pairs in 4 different ways so far. Two more possibilities remain to be considered. At first we express the input voltage V_1 and output current I_2 as a linear transformation of I_1 and V_2 and then we reverse the position. Taking as the starting point Equations 3.35 we manipulate them until V_1 and I_2 appear on the left hand side. Elementary manipulations instead of matrix algebra must be used, since each equation must be rearranged separately. Matrix methods can only be used when complete systems of linear equations are to be handled.

$$V_1 = \frac{B}{D} I_1 + \frac{|A|}{D} V_2'$$

$$-I_2 = I_2' = -\frac{1}{D} I_1 + \frac{C}{D} V_2' \qquad (3.46)$$

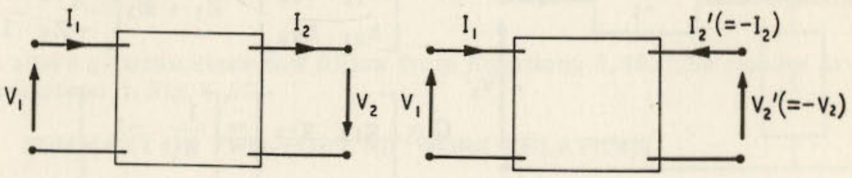
Substituting new symbols with double subscripts for the coefficients in Equations 3.46, the latter can be rewritten in the form

$$\begin{bmatrix} V_1 \\ I_2' \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ V_2' \end{bmatrix}$$

$$V_1 = h_{11} I_1 + h_{12} V_2'$$

$$I_2' = h_{21} I_1 + h_{22} V_2' \qquad (3.47)$$

The assumed direction of the output current has been reversed to conform with conventions usually followed in textbooks on transistors and vacuum valves. The new convention is summarised in Fig. 3.15. The matrix of h-parameters, introduced by Equations 3.47, will be denoted by H and its determinant by $|H|$.



Original sign convention *Fig. 3.15* *Modified sign convention*

The h-parameters may be measured or calculated directly for a two-port network by the method of judicious shortcircuiting and opencircuiting of terminals, used successfully before. The form of Equations 3.47 suggests that the output be shortcircuited to obtain h_{11} and h_{21}, and that the input be opencircuited to obtain h_{12} and h_{22}.

$$h_{11} = \left(\frac{V_1}{I_1}\right)_{V_2'=0}, \quad h_{12} = \left(\frac{V_1}{V_2'}\right)_{I_1=0},$$

$$h_{21} = \left(\frac{I_2'}{I_1}\right)_{V_2'=0}, \quad h_{22} = \left(\frac{I_2'}{V_2'}\right)_{I_1=0} \qquad (3.48)$$

It is found that the above procedure is particularly suited to measurements on transistors, and it is for this reason that transistor characteristics are often stated in terms of h-parameters. The parameter h_{21} is the current gain of a transistor. The foregoing remarks apply to transistors operated under linear conditions only.

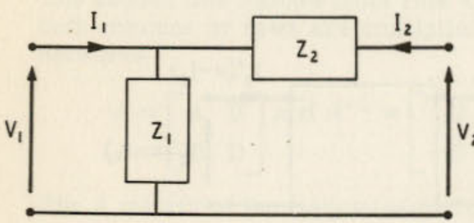
Before considering some examples of h-parameters, it is profitable to observe that $h_{12} = -h_{21}$ for passive bilateral circuits. This follows from the fact, established in the preceding section, that $|A| = 1$ for such circuits.

Let us now find the h-parameters for the L-network of Fig. 3.16. From Equations 3.48 we find

$$h_{11} = \frac{Z_1 Z_2}{Z_1 + Z_2}$$

$$h_{22} = \frac{1}{Z_1 + Z_2}$$

$$h_{12} = \left(\frac{I_2' Z_1}{I_2'(Z_1 + Z_2)}\right)_{I_1=0} = \frac{Z_1}{Z_1 + Z_2} = -h_{21}$$



$$H = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} = \frac{1}{Z_1 + Z_2}\begin{bmatrix} Z_1 Z_2 & Z_1 \\ -Z_2 & 1 \end{bmatrix}$$

$$G = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} \frac{1}{Z_1} & -1 \\ 1 & Z_2 \end{bmatrix}$$

Fig. 3.16

The last line follows because we assume the impedances in Fig. 3.16 to be passive and bilateral.

The sixth and last possible set of four terminal network parameters is obtained on solving Equations 3.47 for the quantities I_1 and V_2'. Since they are usually denoted by the symbols $[g_{ij}] = G$, they are called the *g-parameters* of two-port networks.

$$\begin{bmatrix} I_1 \\ V_2' \end{bmatrix} = \frac{1}{|H|}\begin{bmatrix} h_{22} & -h_{12} \\ -h_{21} & h_{11} \end{bmatrix}\begin{bmatrix} V_1 \\ I_2' \end{bmatrix}$$

$$\begin{bmatrix} I_1 \\ V_2' \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}\begin{bmatrix} V_1 \\ I_2' \end{bmatrix} \qquad (3.49)$$

The opencircuiting and shortcircuiting routine yields the following relations for the g-parameters:

$$g_{11} = \left(\frac{I_1}{V_1}\right)_{I_2'=0}, \quad g_{12} = \left(\frac{I_1}{I_2'}\right)_{V_1=0},$$

$$g_{21} = \left(\frac{V_2'}{V_1}\right)_{I_2'=0}, \quad g_{22} = \left(\frac{V_2'}{V_2'}\right)_{V_1=0}, \qquad (3.50)$$

The off diagonal elements satisfy the relation $g_{12} = -g_{21}$ for passive circuits, as can be seen from their relation to the h-parameters.

As an example the g-parameters for the circuit of Fig. 3.16 are

$$g_{11} = \frac{1}{Z_1},$$

$$g_{22} = Z_2,$$

$$g_{21} = 1 = -g_{12}.$$

It is instructive to check this result against the reciprocal matrix H^{-1} of the same network. The determinant of H is

$$|H| = \frac{Z_1}{Z_1 + Z_2}$$

The above g-parameters now follow from Equations 3.49. The results are summarised in Fig. 3.16

3.6 SUMMARY OF TWO-PORT NETWORK RELATIONS

Four electrical quantities appear at the terminals of a two-port network: V_1, I_1, V_2, I_2. In the preceding sections various ways of expressing two of

these quantities in terms of the others have been introduced. In the present section all possible sets of equations are collected, and relations between the two-port network parameters are tabulated for convenient reference.

Since there are four quantities, we expect six sets of equations, because there are altogether six ways of choosing two objects out of a group of four.

Y-matrix:

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix},$$

Z-matrix:

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix},$$

A-matrix:

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} -V_2 \\ I_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} -V_2 \\ I_2 \end{bmatrix}$$

B-matrix:

$$\begin{bmatrix} -V_2 \\ I_2 \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}$$

H-matrix:

$$\begin{bmatrix} V_1 \\ -I_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ -V_2 \end{bmatrix},$$

G-matrix:

$$\begin{bmatrix} I_1 \\ -V_2 \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ -I_2 \end{bmatrix},$$

The sign convention used in this summary is defined by Fig. 3.17. To avoid the use of negative signs in the column vectors of electrical quantities, primes may be used as shown in the figure.

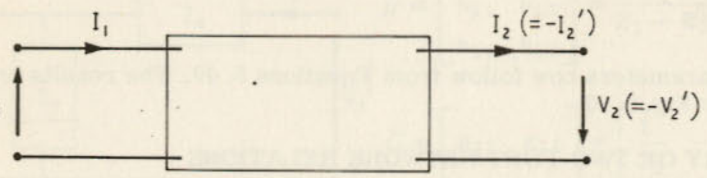


Fig. 3.17

It is useful to express each matrix in terms of all the others and to tabulate the results as in Table 3.1. Some of the matrices are reciprocals of others-this is also indicated in the table. To acquire familiarity the reader should work through the table, checking systematically all the entries.

The determinants of two-port network matrices enter into many calculations. For easy reference Table 3.2 gives the determinants in terms of the elements of the remaining matrices.

The foregoing summary applies to all *linear* networks without any additional restrictions. The matrices possess a number of special properties mainly symmetry properties, in the case of circuits composed exclusively of *passive bilateral* impedance elements. The special properties are listed below.

1. The admittance matrix is symmetric (see p. 72)

$$Y^t = Y \text{ or } y_{ij} = y_{ji}.$$

2. The impedance matrix is symmetric (see p. 77)

$$Z^t = Z \text{ or } z_{ij} = z_{ji}.$$

3. The determinant of the A-matrix and its reciprocal equals unity (see p. 82)

$$|A| = |A^{-1}| = |B| = 1.$$

4. For the H- and G-matrices it was established that

$$h_{12} = -h_{21}, \quad g_{12} = -g_{21}.$$

The last property should not be confused with antisymmetry of square matrices. On reference to page 43 it will be recalled that the diagonal elements of symmetric matrices are zero, which is not the case with H- and G-matrices.

The above list of special properties should be used to simplify the relations tabulated in Tables 3.1 and 3.2, whenever these are applied to passive bilateral networks.

Further restrictions on the generality of networks result in additional simplication of their matrices. *Symmetric two-port networks* are a case worth noting.

1. The diagonal elements of the Y, Z, A, B matrices are equal (see p. 77)

$$y_{11} = y_{22}, \; z_{11} = z_{22}, \; A = D, \; b_{11} = b_{22}.$$

2. The determinants of the H and G matrices are equal to unity

$$|H| = 1, \quad |G| = 1.$$

Before the subject of two-port network parameters is closed, a word of comment on sign conventions is called for. In the present treatment the matrices have been introduced using a sign convention that follows naturally from the basic mesh equations of a general circuit. The convention was then changed by stages to follow accepted practice in connection with the general circuit constants and the h- and g-parameters. The y- and z-parameters are sometimes defined in relation to sign conventions other than the one used above. The

Table 3.1

	In terms of		
	Y	*Z*	*A*
Y	$\begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix}$	$Z^{-1} = \frac{1}{Z}\begin{bmatrix} z_{22} & -z_{12} \\ -z_{21} & z_{11} \end{bmatrix}$	$\frac{1}{B}\begin{bmatrix} D & \lvert A\rvert \\ 1 & A \end{bmatrix}$
Z	Y^{-1} $\frac{1}{\lvert Z\rvert}\begin{bmatrix} y_{22} & -y_{12} \\ -y_{21} & y_{11} \end{bmatrix}$	$\begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix}$	$\frac{1}{C}\begin{bmatrix} A & -\lvert A\rvert \\ -1 & D \end{bmatrix}$
A	$\frac{1}{y_{21}}\begin{bmatrix} y_{22} & 1 \\ \lvert Y\rvert & y_{11} \end{bmatrix}$	$\frac{-1}{z_{21}}\begin{bmatrix} z_{11} & \lvert Z\rvert \\ 1 & z_{22} \end{bmatrix}$	$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$
B	$\begin{bmatrix} y_{11} & -1 \\ -\lvert Y\rvert & y_{22} \end{bmatrix}$	$\frac{1}{z_{12}}\begin{bmatrix} -z_{22} & \lvert Z\rvert \\ 1 & -z_{11} \end{bmatrix}$	A^{-1} $\frac{1}{\lvert A\rvert}\begin{bmatrix} D & -B \\ -C & A \end{bmatrix}$
H	$\frac{1}{y_{11}}\begin{bmatrix} 1 & y_{12} \\ -y_{21} & \lvert Y\rvert \end{bmatrix}$	$\frac{1}{z_{22}}\begin{bmatrix} \lvert Z\rvert & -z_{12} \\ z_{21} & 1 \end{bmatrix}$	$\frac{1}{D}\begin{bmatrix} B & \lvert A\rvert \\ -1 & C \end{bmatrix}$
G	$\frac{1}{y_{22}}\begin{bmatrix} \lvert Y\rvert & -y_{12} \\ y_{21} & 1 \end{bmatrix}$	$\frac{1}{z_{11}}\begin{bmatrix} 1 & z_{12} \\ -z_{21} & \lvert Z\rvert \end{bmatrix}$	$\frac{1}{A}\begin{bmatrix} C & -\lvert A\rvert \\ 1 & B \end{bmatrix}$

nett effect of such changes is to alter the sign of some of the matrix elements. To avoid confusion the reader is advised to watch for these differences when reading other books or papers.

Whenever general circuit constants or mixed parameters are used in practical problems, it may prove tiresome to have to write primes repeatedly above the output voltage or current. They can be omitted, of course, provided it is remembered what sign convention is actually used.

Table 3.1 (continuation)

	In terms of		
	B	*H*	*G*
Y	$\frac{-1}{b_{12}}\begin{bmatrix} b_{11} & 1 \\ \lvert B\rvert & b_{22} \end{bmatrix}$	$\frac{1}{h_{11}}\begin{bmatrix} 1 & h_{12} \\ -h_{21} & \lvert H\rvert \end{bmatrix}$	$\frac{1}{g_{22}}\begin{bmatrix} \lvert G\rvert & -g_{12} \\ g_{21} & 1 \end{bmatrix}$
Z	$\frac{1}{b_{21}}\begin{bmatrix} -b_{22} & 1 \\ \lvert B\rvert & -b_{11} \end{bmatrix}$	$\frac{1}{h_{22}}\begin{bmatrix} \lvert H\rvert & -h_{12} \\ h_{21} & 1 \end{bmatrix}$	$\frac{1}{g_{11}}\begin{bmatrix} 1 & g_{12} \\ -g_{21} & \lvert G\rvert \end{bmatrix}$
A	B^{-1} $\frac{1}{\lvert B\rvert}\begin{bmatrix} b_{22} & -b_{12} \\ -b_{21} & b_{11} \end{bmatrix}$	$\frac{-1}{h_{21}}\begin{bmatrix} \lvert H\rvert & h_{11} \\ h_{22} & 1 \end{bmatrix}$	$\frac{1}{g_{21}}\begin{bmatrix} 1 & g_{22} \\ g_{11} & \lvert G\rvert \end{bmatrix}$
B	$\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$	$\frac{1}{h_{12}}\begin{bmatrix} 1 & -h_{11} \\ -h_{22} & \lvert H\rvert \end{bmatrix}$	$\frac{-1}{g_{12}}\begin{bmatrix} \lvert G\rvert & -g_{22} \\ -g_{11} & 1 \end{bmatrix}$
H	$\frac{1}{b_{11}}\begin{bmatrix} -b_{12} & 1 \\ -\lvert B\rvert & -b_{21} \end{bmatrix}$	$\begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}$	$H^{-1} = \frac{1}{\lvert G\rvert}\begin{bmatrix} g_{22} & -g_{12} \\ -g_{21} & g_{11} \end{bmatrix}$
G	$\frac{1}{b_{22}}\begin{bmatrix} -b_{21} & -1 \\ \lvert B\rvert & -b_{12} \end{bmatrix}$	$G^{-1} = \frac{1}{\lvert H\rvert}\begin{bmatrix} h_{22} & -h_{12} \\ -h_{21} & h_{11} \end{bmatrix}$	$\begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$

Table 3.2

	In terms of					
	Y	Z	A	B	H	G
$\lvert Y\rvert$		$\dfrac{1}{\lvert Z\rvert}$	$\dfrac{C}{B}$	$\dfrac{b_{21}}{b_{12}}$	$\dfrac{h_{22}}{h_{11}}$	$\dfrac{g_{11}}{g_{22}}$
$\lvert Z\rvert$	$\dfrac{1}{\lvert Y\rvert}$		$\dfrac{B}{C}$	$\dfrac{b_{12}}{b_{21}}$	$\dfrac{h_{11}}{h_{22}}$	$\dfrac{g_{22}}{g_{11}}$
$\lvert A\rvert$	$\dfrac{y_{12}}{y_{21}}$	$\dfrac{z_{12}}{z_{21}}$		$\dfrac{1}{\lvert B\rvert}$	$-\dfrac{h_{12}}{h_{21}}$	$-\dfrac{g_{12}}{g_{21}}$
$\lvert F\rvert$	$\dfrac{y_{21}}{y_{12}}$	$\dfrac{z_{21}}{z_{12}}$	$\dfrac{1}{\lvert A\rvert}$		$-\dfrac{h_{21}}{h_{12}}$	$-\dfrac{g_{21}}{g_{12}}$
$\lvert H\rvert$	$\dfrac{y_{22}}{y_{11}}$	$\dfrac{z_{11}}{z_{22}}$	$\dfrac{A}{D}$	$\dfrac{b_{22}}{b_{11}}$		$\dfrac{1}{\lvert G\rvert}$
$\lvert G\rvert$	$\dfrac{y_{11}}{y_{22}}$	$\dfrac{z_{22}}{z_{11}}$	$\dfrac{D}{A}$	$\dfrac{b_{11}}{b_{22}}$	$\dfrac{1}{\lvert H\rvert}$	

3.7 CASCADE CONNECTIONS OF TWO-PORT NETWORKS

The advantages of the matrix form of two-port network equations become apparent when two or more networks are interconnected. There are 5 ways of connecting two-port networks:

1. cascade,
2. parallel,
3. series,
4. series-parallel, } mixed connections.
5. parallel-series, }

Each connection is most conveniently handled by one of the two-port network matrices, as will be explained in this and the following two sections.

Starting with the cascade connection, let us write down the equations of the two separate networks shown in Fig. 3.18, which is a reproduction of Fig. 1.3 of Chapter 1.

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A^{(1)} & B^{(1)} \\ C^{(1)} & D^{(1)} \end{bmatrix} \begin{bmatrix} V_2' \\ I_2 \end{bmatrix}, \qquad W_1 = A^{(1)}W_2 \tag{3.51}$$

$$\begin{bmatrix} V_2' \\ I_2 \end{bmatrix} = \begin{bmatrix} A^{(2)} & B^{(2)} \\ C^{(2)} & D^{(2)} \end{bmatrix} \begin{bmatrix} V_3' \\ I_3 \end{bmatrix}, \qquad W_2 = A^{(2)}W_3 \tag{3.52}$$

We observe that the column vector on the left hand side of Equations 3.52 is identical with the column vector on the right hand side of Equations 3.51. This is inherent in the method of connection and is made obvious by the assumed polarity of the terminal voltages and currents. Substituting Equations 3.52 into Equations 3.51 we obtain

$$W_1 = A^{(1)}A^{(2)}W_3 = AW_3 \tag{3.53}$$

Equation 3.53 relates the input and output of the combined two-port network of Fig. 3.18. The A-matrix of the combined network is the product of the A-matrices of the individual networks.

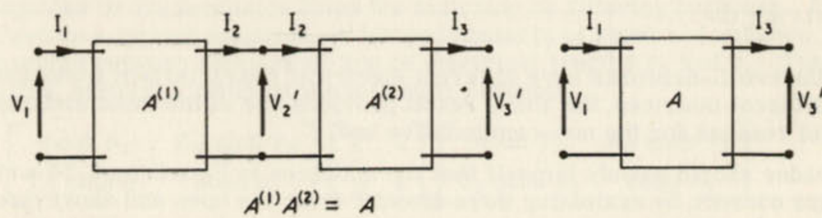


Fig. 3.18

As an application of the foregoing result let us work out the A-matrix of the L-network of Fig. 3. 19, which is the same as the one shown in Fig. 3. 10, except for the marking of the output voltage and current. The latter is now chosen to agree with the symbols of Fig. 3. 18. The L-network is a cascade

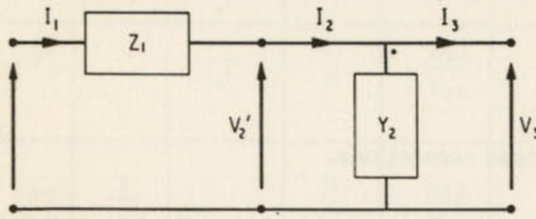


Fig. 3. 19

connection of the series impedance Z_1 and the shunt admittance Y_2, whose A-matrices have been found in the section on general circuit constants, Equations 3. 38 and 3. 39.

$$A^{(1)} = \begin{bmatrix} 1 & Z_1 \\ 0 & 1 \end{bmatrix}, \qquad A^{(2)} = \begin{bmatrix} 1 & 0 \\ Y_2 & 1 \end{bmatrix}$$

$$L^{(1)} = A^{(1)}A^{(2)} = \begin{bmatrix} 1 & Z_1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ Y_2 & 1 \end{bmatrix} = \begin{bmatrix} 1 + Z_1 Y_2 & Z_1 \\ Y_2 & 1 \end{bmatrix} \tag{3. 54}$$

where $L^{(1)}$ is the matrix of the circuit shown in Fig. 3. 19.

Reversing the sequence of the shunt and series elements of Fig. 3. 19 we obtain another L-network, this time like the one of Fig. 3. 16. Its matrix is found by multiplying $A^{(1)}$ and $A^{(2)}$ in reverse order. Denoting the product by $L^{(2)}$ we obtain

$$L^{(2)} = A^{(2)}A^{(1)} = \begin{bmatrix} 1 & 0 \\ Y_2 & 1 \end{bmatrix}\begin{bmatrix} 1 & Z_1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & Z_1 \\ Y_2 & Z_1 Y_2 + 1 \end{bmatrix} \tag{3. 55}$$

As could be expected from the non-commutative law of matrix multiplication

$$L^{(1)} \neq L^{(2)}.$$

Since the two L-networks have different electrical characteristics and, therefore, different matrices, the above result provides one of the most striking physical reasons for the non-commutative law.

The reader should satisfy himself that the matrices in Equations 3. 54 and 3. 55 are correct, by evaluating them directly from the open and short circuit conditions for the general circuit constants (Section 3. 4).

The process of cascading two-port networks applies to any number of networks beyond two. In every case the matrix of general circuit constants for the resulting network is the product of individual matrices.

As an example of this procedure we build up a T-network from a series impedance followed by a shunt admittance and followed again by a series impedance. The A-matrix of the T-network, found in this way, is

$$A = \begin{bmatrix} 1 & Z_1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ Y_2 & 1 \end{bmatrix}\begin{bmatrix} 1 & Z_3 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 + Z_1 Y_2 & Z_1 \\ Y_2 & 1 \end{bmatrix}\begin{bmatrix} 1 & Z_3 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 + Z_1 Y_2 & Z_1 + Z_3(1 + Z_1 Y_2) \\ Y_2 & Y_2 Z_3 + 1 \end{bmatrix} \tag{3. 56}$$

A π-network may be similarly built up from a series element flanked by two shunt elements. Its A-matrix is

$$\begin{bmatrix} 1 & 0 \\ Y_1 & 1 \end{bmatrix}\begin{bmatrix} 1 & Z_2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ Y_3 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 \\ Y_1 & 1 \end{bmatrix}\begin{bmatrix} 1 + Z_2 Y_3 & Z_2 \\ Y_3 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 + Z_2 Y_3 & Z_2 \\ Y_3 + Y_1(1 + Z_2 Y_3) & Y_1 Z_2 + 1 \end{bmatrix} \tag{3. 57}$$

In working out the Equations 3. 56 and 3. 57, Equation 3. 54 was used to save labour. The circuit elements are labelled as in Figs. 3. 9 and 3. 4, except that shunt elements are written as admittances and series elements as impedances to avoid fractional expressions.

The foregoing methods prove useful when dealing with obstacles inserted in waveguides or transmission lines for matching or filtering purposes. Fig.3.20 represents a typical situation. A lumped obstacle of shunt susceptance jy is inserted between sections of line of electrical lengths θ_1 and θ_2. The resulting matrix is given by the following product:

$$\begin{bmatrix} \cosh\theta_1 & Z_0 \sinh\theta_1 \\ Y_0 \sinh\theta_1 & \cosh\theta_1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ jy & 1 \end{bmatrix}\begin{bmatrix} \cosh\theta_2 & Z_0 \sinh\theta_2 \\ Y_0 \sinh\theta_2 & \cosh\theta_2 \end{bmatrix}$$

Multiplication and simplification yields the single matrix

$$\begin{bmatrix} \cosh(\theta_1+\theta_2)+jyZ_0\sinh\theta_1\cosh\theta_2 & Z_0[\sinh(\theta_1+\theta_2)+jyZ_0\sinh\theta_1\sinh\theta_2] \\ Y_0\sinh(\theta_1+\theta_2)+jy\cosh\theta_1\cosh\theta_2 & \cosh(\theta_1+\theta_2)+jyZ_0\cosh\theta_1\sinh\theta_2 \end{bmatrix} \tag{3.58}$$

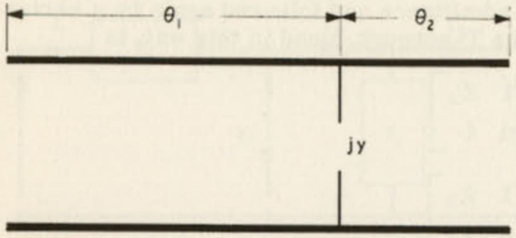


Fig. 3.20

The method of cascading two-port networks may be applied to find the properties of a *circuit terminated* in a load impedance Z_l and driven by a generator having an internal impedance Z_g. The arrangement is shown in Fig. 3.21. The generator impedance and the load are considered to be separate two-port networks. In what follows the matrix of the circuit under consideration will be written without any identification superscript, e.g. A or Y, while possible cascade connections will be marked by superscripts, as shown in the figure.

First let us find the voltage ratio V_2'/V_1. Since the load is effectively a shunt admittance with its output open circuited, this ratio is equal to the reciprocal of the element $A^{(1)}$ of the matrix $A^{(1)}$ (see page 79). Now

$$A^{(1)} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} 1 & 0 \\ \frac{1}{Z_l} & 1 \end{bmatrix} = \begin{bmatrix} A+\frac{B}{Z_l} & B \\ C+\frac{D}{Z_l} & D \end{bmatrix} \tag{3.59}$$

and

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A+\frac{B}{Z_l} & B \\ C+\frac{D}{Z_l} & D \end{bmatrix}\begin{bmatrix} V_2' \\ I_0 \end{bmatrix}$$

where $I_0 = 0$ permanently.

Hence $$V_2'/V_1 = \frac{Z_l}{B+AZ_l} \tag{3.60}$$

It is convenient to have this ratio expressed in terms of other sets of two-port network parameters, besides the general circuit constants. This is done

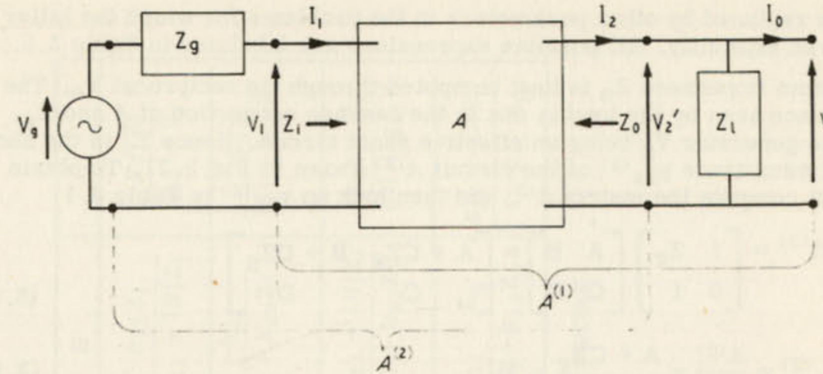


Fig. 3.21

by substituting the appropriate expressions for A and B from Table 3.1. All possible results are listed in Table 3.3 for reference.

Next we compute the current ratio I_2/I_1 from the relation

$$I_2 = -\frac{V_2'}{Z_l}$$

where V_2 is related to I_1 by the element $C^{(1)}$ of the matrix of Equation 3.59. Hence

$$\frac{I_2}{I_1} = \frac{1}{D + CZ_l} \tag{3.61}$$

Again this ratio is expressed in terms of all the other two-port network parameters and tabulated in Table 3.3.

In practical problems it is often necessary to know the *input and output impedance of a terminated network*, as seen by the generator and load respectively. The input impedance Z_i is effectively the open circuit input impedance (see page 74) of the cascade connection $A^{(1)}$. It is the element $z_{11}^{(1)}$ of the impedance matrix $Z^{(1)}$ Since the matrix $A^{(1)}$ of this connection is given by Equation 3.59, the impedance parameter $z_{11}^{(1)}$ is easily found with the help of Table 3.1.

$$z_{11}^{(1)} = \frac{A^{(1)}}{C^{(1)}} = \frac{B + AZ_l}{D + CZ_l} = Z_i \tag{3.62}$$

Equation 3.62 gives the input impedance of the terminated network in terms of its general circuit constants and the load Z_l. The general circuit constants

may be replaced by other parameters in the problems for which the latter are given explicitly. All possible expressions are tabulated in Table 3.3.

The output impedance Z_0 is best computed through its reciprocal Y_0. The admittance seen by the load is due to the cascade connection of A and Z_g, with the generator V_g being an effective short circuit. Hence Y_0 is the short circuit admittance $y_{22}{}^{(2)}$ of the circuit $A^{(2)}$ shown in Fig. 3.21. To obtain it we first compute the matrix $A^{(2)}$ and then look up $y_{22}{}^{(2)}$ in Table 3.1

$$A^{(2)} = \begin{bmatrix} 1 & Z_g \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A + CZ_g & B + DZ_g \\ C & D \end{bmatrix} \tag{3.63}$$

$$y_{22}{}^{(2)} = \frac{A^{(2)}}{B^{(2)}} = \frac{A + CZ_g}{B + DZ_g} = Y_0 \tag{3.64}$$

The reciprocal of Equation 3.64, expressed in terms of all the parameters of the network A, is listed in Table 3.3.

As an example of the foregoing formulae we consider an ideal transformer feeding power from a generator to a load, as shown in Fig. 3.22. By Equation 3.60 and Fig. 3.13 the input and output voltage ratio is equal to the turns ratio.

$$\frac{V_2'}{V_1} = n \tag{3.65}$$

Equation 3.61 checks that the current ratio at the terminals is the reciprocal of the turns ratio.

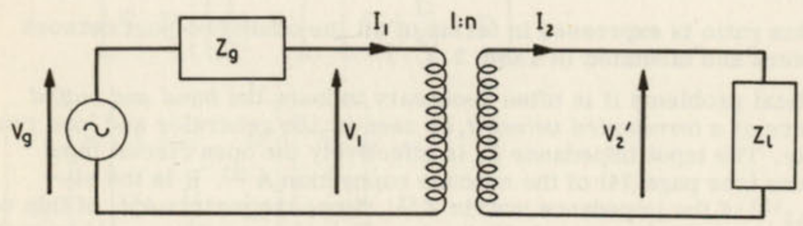


Fig. 3.22

The input impedance of the transformer on load, as given by Equation 3.62, is

$$Z_i = \frac{1}{n^2} Z_1 \tag{3.66}$$

	Y	Z	A	B	H	G
$\frac{V_2'}{V_1}$	$\frac{y_{21}Z_l}{1 + y_{22}Z_l}$	$\frac{-z_{21}Z_l}{\lvert Z\rvert + z_{11}Z_l}$	$\frac{Z_l}{B + AZ_l}$	$\frac{\lvert B\rvert Z_l}{-b_{12} + b_{22}Z_l}$	$\frac{-h_{21}Z_l}{h_{11} + \lvert H\rvert Z_l}$	$\frac{g_{21}Z_l}{g_{22} + Z_l}$
$\frac{I_2}{I_1}$	$\frac{y_{21}}{y_{11} + \lvert Y\rvert Z_l}$	$\frac{-z_{21}}{z_{22} + Z_l}$	$\frac{1}{D + CZ_l}$	$\frac{\lvert B\rvert}{b_{11} - b_{21}Z_l}$	$\frac{-h_{21}}{1 + h_{22}Z_l}$	$\frac{g_{21}}{\lvert G\rvert + h_{22}Z_l}$
Z_i	$\frac{1 + y_{22}Z_l}{y_{11} + \lvert Y\rvert Z_l}$	$\frac{\lvert Z\rvert + z_{11}Z_l}{z_{22} + Z_l}$	$\frac{B + AZ_l}{D + CZ_l}$	$\frac{-b_{12} + b_{22}Z_l}{b_{11} - b_{21}Z_l}$	$\frac{h_{11} + \lvert H\rvert Z_l}{1 + h_{22}Z_l}$	$\frac{g_{22} + Z_l}{\lvert G\rvert + g_{11}Z_l}$
Z_o	$\frac{1 + y_{11}Z_g}{y_{22} + \lvert Y\rvert Z_g}$	$\frac{\lvert Z\rvert + z_{22}Z_g}{z_{11} + Z_g}$	$\frac{B + DZ_g}{A + CZ_g}$	$\frac{-b_{12} + b_{11}Z_g}{b_{22} - b_{21}Z_g}$	$\frac{h_{11} + Z_g}{\lvert H\rvert + h_{22}Z_g}$	$\frac{g_{22} + \lvert G\rvert Z_g}{1 + g_{11}Z_g}$

TABLE 3.3. PROPERTIES OF THE TERMINATED TWO-PORT NETWORK
(see Fig. 3.21 for symbols and sign convention.)

while its output impedance, as given by the reciprocal of Equation 3.64, is

$$Z_0 = n^2 Z_g \tag{3.67}$$

The above properties of the ideal transformer are well known. They are usually derived without recourse to matrix algebra, but the reader will, no doubt, appreciate the conciseness and lucidity of matrix methods on this example.

3.8 PARALLEL AND SERIES CONNECTIONS OF TWO-PORT NETWORKS

Having dealt with the problem of cascading two-port networks we pass on to the *parallel connection*. Two four terminal networks are said to be connected in parallel if their corresponding terminals are strapped as shown in Fig. 3.23 The matrices of the individual networks, $Y^{(1)}$ and $Y^{(2)}$, are assumed to be known, and the problem is to find the matrix Y of the combined network.

Observing that the component networks have the same terminal voltages but different currents, we write their equations in terms of the y-parameters and add them.

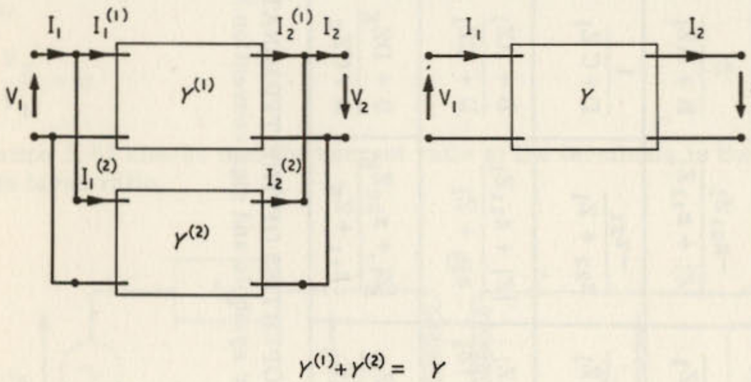


Fig. 3.23

$$I^{(1)} = Y^{(1)} V$$

$$I^{(2)} = Y^{(2)} V \tag{3.68}$$

$$I^{(1)} + I^{(2)} = I = \left(Y^{(1)} + Y^{(2)}\right)V = YV \tag{3.69}$$

The symbols used in Equations 3.68 and 3.69 are defined by Fig. 3.23. Equation 3.69 is obtained by the rule for the addition of matrices, together with the distributive law of matrix algebra (Section 2.8 and 2.9).

From Equation 3.69 we conclude that the admittance matrix of a parallel connection of two-port networks is the sum of the matrices of individual networks. Although Equation 3.69 establishes this rule for two networks only, an extension to any number of networks is obvious. The reader should not fail to note the close analogy between matrix relations for two-port networks on the one hand, and equations of ordinary admittances in parallel on the other, Fig. 1.10, makes the similarity clear.

A typical example of networks in parallel is provided by the bridged T-circuit, which can be assembled from a T-network and a series impedance, as shown in Fig. 3.24.

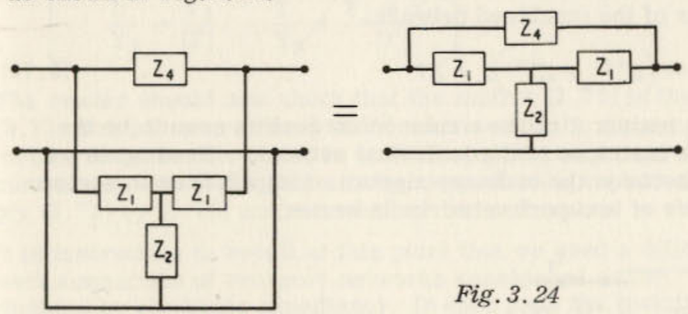


Fig. 3.24

The Y-matrix of the series element has been found earlier (Fig. 3.7), but the admittance matrix of the T-network must be calculated now. Starting from (3.31) we can obtain the reciprocal of the Z-matrix given there. This is

$$\begin{bmatrix} \dfrac{Z_1 + Z_2}{|Z|} & \dfrac{Z_2}{|Z|} \\[2ex] \dfrac{Z_2}{|Z|} & \dfrac{Z_1 + Z_2}{|Z|} \end{bmatrix} \tag{3.70}$$

where $|Z| = Z_1(Z_1 + 2Z_2)$ is the determinant of the symmetric T-network. Addition of the above matrix to the matrix of Z_4 yields the admittance matrix of the bridged T.

$$\begin{bmatrix} \dfrac{Z_1 + Z_2}{|Z|} + \dfrac{1}{Z_4} & \dfrac{Z_2}{|Z|} + \dfrac{1}{Z_4} \\[2ex] \dfrac{Z_2}{|Z|} + \dfrac{1}{Z_4} & \dfrac{Z_1 + Z_2}{|Z|} + \dfrac{1}{Z_4} \end{bmatrix} \tag{3.71}$$

The next on our list of two-port network connections is the *series connection* shown in Fig. 3.25. Writing down the equations of the component networks in terms of their impedance matrices we find

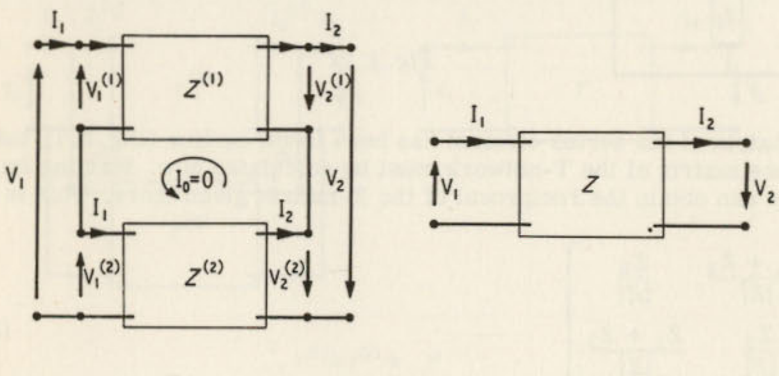
$$V^{(1)} = Z^{(1)} I$$
$$V^{(2)} = Z^{(2)} I \qquad (3.72)$$

The foregoing equations can only be written down on the assumption that the currents in the corresponding ports of the individual circuits are equal, as shown in Fig. 3.25. The assumption is expressed by the condition $I_0 = 0$, marked in the figure, where I_0 is a loop current that might conceivably be set up in the extra loop created by the series connection.

Addition of Equations 3.72 yields the required relation between the terminal voltages and currents of the combined network.

$$V^{(1)} + V^{(2)} = V = \left(Z^{(1)} + Z^{(2)}\right) I = ZI \qquad (3.73)$$

Hence the impedance matrix Z of the combined network is seen to be the sum of the impedance matrices of the individual networks. Here again we find a close analogy between the ordinary algebra of impedances in series and the matrix algebra of two-port networks in series.



$Z^{(1)} + Z^{(2)} = Z$

Fig. 3.25

An example of series connected networks is again provided by the bridged T, which can be decomposed into a shunt impedance and a π-network, as shown in Fig. 3.26. The impedance matrix of the π-network must be calculated by taking the reciprocal of (3.24), while the matrix of the shunt element is given in Fig. 3.11.

The Z-matrix of the Π-network is:

$$\begin{bmatrix} \dfrac{Y_1 + Y_4}{|Y|} & -\dfrac{Y_4}{|Y|} \\[2ex] -\dfrac{Y_4}{|Y|} & \dfrac{Y_1 + Y_4}{|Y|} \end{bmatrix} \qquad (3.74)$$

where $|Y| = Y_1 (Y_1 + 2Y_4)$.
Z-matrix of bridged T-network:

$$\begin{bmatrix} \dfrac{1}{Y_2} + \dfrac{Y_1 + Y_4}{|Y|} & \dfrac{1}{Y_2} - \dfrac{Y_4}{|Y|} \\[2ex] \dfrac{1}{Y_2} - \dfrac{Y_4}{|Y|} & \dfrac{1}{Y_2} + \dfrac{Y_1 + Y_4}{|Y|} \end{bmatrix} \qquad (3.75)$$

The reader should now check that the matrix (3.75) is the reciprocal of (3.71) since they are the Z and Y matrices of the same network. An obvious method to make this check is to evaluate the reciprocal of, say, (3.71) and compare it with (3.75), but a better way of using matrix algebra is to multiply (3.71) by (3.75) and see whether the unit matrix is obtained.

It is instructive to recall at this point that we used a different matrix for each connection of two-port networks considered so far. This procedure was dictated by algebraic expediency. In each case the matrix to use was suggested by the electrical quantities common to the networks. Thus, in the parallel connection the terminal voltages were common to the two networks and the Y-matrix made it possible to utilise this fact. In the cascade connection the output quantities of one network were equal to the input quantities of the other and the A-matrix proved naturally suitable.

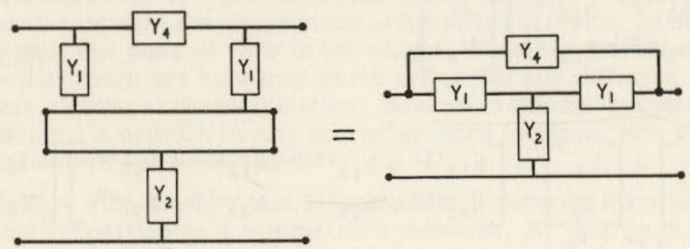
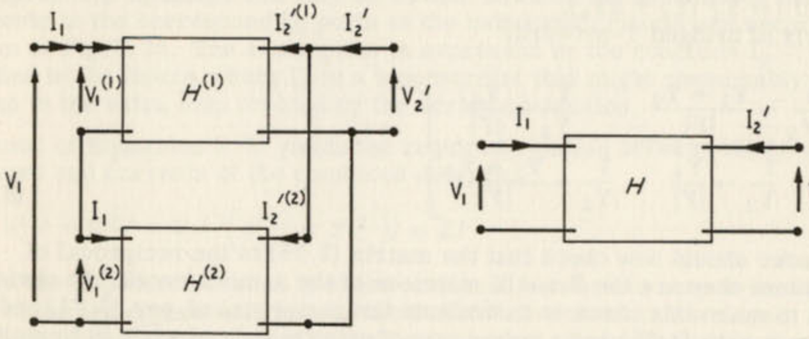


Fig. 3.26

The same approach will now be followed in the discussion of the remaining, mixed, connections. It will be found that the mixed parameters are convenient in the solution of this problem.

3.9 MIXED CONNECTIONS OF TWO-PORT NETWORKS

The parallel and series connections of two-port networks lend themselves to be mixed together in two ways. First the input ports of two networks are connected serieswise, while the output ports are strapped in parallel (Fig. 3.27). Next the situation is reversed (Fig. 3.28). Let us solve the former problem first and obtain the matrix of the combined network in terms of the matrices of the individual networks.



$H^{(1)}+H^{(2)}=H$

Fig. 3.27

The problem is easily tackled in terms of the H-matrices. The equations of the individual networks are

$$\begin{bmatrix} V_1^{(1)} \\ I_2'^{(1)} \end{bmatrix} = \begin{bmatrix} h_{11}^{(1)} & h_{12}^{(1)} \\ h_{21}^{(1)} & h_{22}^{(1)} \end{bmatrix} \begin{bmatrix} I_1 \\ V_2' \end{bmatrix}$$

$$\begin{bmatrix} V_1^{(2)} \\ I_2'^{(2)} \end{bmatrix} = \begin{bmatrix} h_{11}^{(2)} & h_{12}^{(2)} \\ h_{21}^{(2)} & h_{22}^{(2)} \end{bmatrix} \begin{bmatrix} I_1 \\ V_2' \end{bmatrix} \tag{3.76}$$

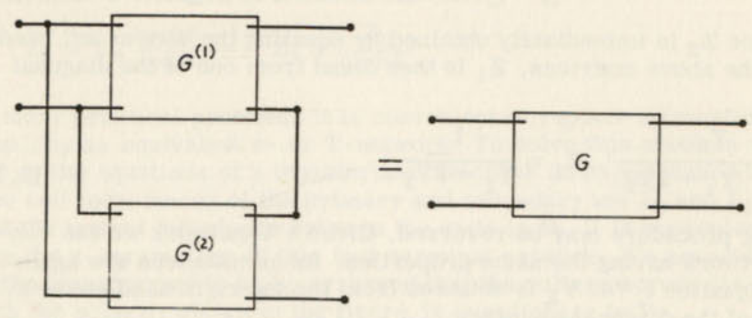
Addition of Equations 3.76 yields

$$\begin{bmatrix} V_1^{(1)} \\ I_2'^{(1)} \end{bmatrix} + \begin{bmatrix} V_1^{(2)} \\ I_2'^{(2)} \end{bmatrix} = \begin{bmatrix} V_1 \\ I_2' \end{bmatrix} = \begin{bmatrix} h_{11}^{(1)} + h_{11}^{(2)} & h_{12}^{(1)} + h_{12}^{(2)} \\ h_{21}^{(1)} + h_{21}^{(2)} & h_{22}^{(1)} + h_{22}^{(2)} \end{bmatrix} \begin{bmatrix} I_1 \\ V_2' \end{bmatrix}$$

$$\begin{bmatrix} V_1 \\ I_2' \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ V_2' \end{bmatrix} \tag{3.77}$$

Hence the H-matrix of the series-parallel connection is the sum of the H-matrices of the component networks as summarised in Fig. 3.27.

The parallel-series connection of Fig. 3.28 is best treated in terms of the g-parameters. The solution follows exactly the lines of Equations 3.76 and 3.77, with the result that the G-matrix of the combined network is given by the sum of the G-matrices of the individual networks, as summarised in Fig. 3.28.



$G^{(1)}+G^{(2)}= G$

Fig. 3.28

With the above problem all possibilities of interconnecting two-port networks are exhausted. We have seen that each connection is best tackled in terms of a particular matrix, which means that whenever other matrices are required in practical problems, they must be found with the help of Table 3.1.

3.10 EQUIVALENT CIRCUITS

The point of view adopted in the preceding sections of this chapter was that of circuit analysis. The typical problem of analysis is to find the electrical characteristics of a given circuit. The characteristics assume the form of parameters for two-port networks, and it has been our task so far to learn how to calculate the parameters of specified circuits. In the present section we shift our point of view in two stages. First we consider the question whether there are networks which are physically different, and yet have the same electrical characteristics. Next we tackle the problem of how to assemble a network having characteristics identical with a 'black box' whose parameters are prescribed.

To start with we show that a symmetric T-network may have exactly the same properties as a symmetric π-network. By this we mean that it is possible to assemble a T-network by judiciously choosing impedances to have, say, the same A-matrix as the given π-network. To demonstrate this we write down the A-matrix of the T-network and equate it to the A-matrix of the π-network. We start with the Y- and Z- matrices as given by Equations 3.24 and 3.31, and obtain the corresponding A-matrices with the help of Table 3.1

$$\begin{bmatrix} 1+\dfrac{Z_1}{Z_2} & \dfrac{Z_1(Z_1+2Z_2)}{Z_2} \\ \dfrac{1}{Z_2} & 1+\dfrac{Z_1}{Z_2} \end{bmatrix} = \begin{bmatrix} 1+\dfrac{Y_1}{Y_2} & \dfrac{1}{Y_2} \\ \dfrac{Y_1(Y_1+2Y_2)}{Y_2} & 1+\dfrac{Y_1}{Y_2} \end{bmatrix} \tag{3.78}$$

The impedance Z_2 is immediately obtained by equating the bottom left hand elements of the above matrices. Z_1 is then found from one of the diagonal elements.

$$Z_2 = \frac{Y_2}{Y_1\,(Y_1+2Y_2)},\quad Z_1 = \frac{1}{Y_1+2Y_2} \tag{3.79}$$

The foregoing procedure may be reversed. Given a T-network we can construct a π-network having the same properties. Its admittances are again found from Equation 3.78. Y_2 is obtained from the top right hand element, and Y_1 from one of the diagonal elements.

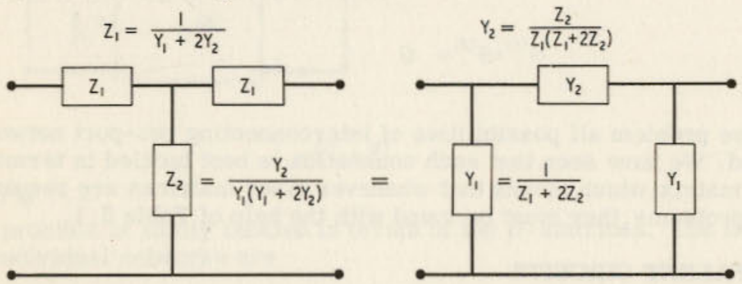


Fig. 3.29

$$Y_2 = \frac{Z_2}{Z_1\,(Z_1+2Z_2)},\quad Y_1 = \frac{1}{Z_1+2Z_2} \tag{3.79}$$

The results are summarised in Fig. 3.29.

The reader is invited to apply the open and short circuiting technique to the circuits of Fig. 3.29, to verify that their A-matrices are identical and equal to those in Equation 3.78. They are, therefore, equivalent networks.

By the same procedure a variety of network configurations can be set up having characteristics identical with a given network.

A useful example is provided by a length of transmission line and its equivalent π- and T-networks. The general circuit constants of a section of line of electrical length θ are given in Equation 3.43. It only remains to equate their matrix to Equation 3.78 to obtain the equivalent circuits.

Equivalent π-network of transmission line:

$$Y_1 = \frac{\cosh\theta - 1}{Z_0\,\sinh\theta},\quad Y_2 = \frac{1}{Z_0\,\sinh\theta} \tag{3.80}$$

Equivalent T-network of transmission line:

$$Z_1 = Z_0\,\frac{\cosh\theta - 1}{\sinh\theta},\quad Z_2 = \frac{Z_0}{\sinh\theta} \tag{3.81}$$

In many practical problems it is convenient to replace a transformer (not ideal) by an equivalent π- or T-network. To solve this problem we must first set up the equations of a transformer which we do by reference to Fig. 3.30. The self inductances of the primary and secondary are L_1 and L_2 respectively, and the mutual inductance between the coils is M. It is convenient to calculate the z-parameters of this four terminal network. We assume the windings of the transformer to be so arranged that the voltage across the secondary, with the polarity shown in the figure, is in antiphase to V_1.

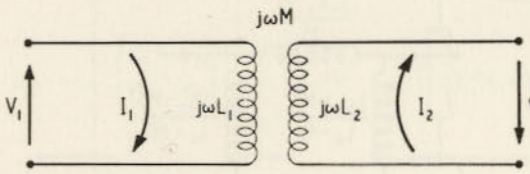


Fig. 3.30

Open circuiting the secondary we find

$$z_{11} = \left(\frac{V_1}{I_1}\right)_{I_2=0} = j\omega L_1$$

$$z_{21} = \left(\frac{V_2}{I_1}\right)_{I_2=0} = -j\omega M$$

The above expressions show the primary and secondary voltages in antiphase, as required. Next we open circuit the primary to obtain the remaining parameters.

$$z_{12} = \left(\frac{V_1}{I_2}\right)_{I_1=0} = -j\omega M$$

$$z_{22} = \left(\frac{V_2}{I_2}\right)_{I_1=0} = j\omega L_2$$

Hence the Z-matrix of the transformer is

$$\begin{bmatrix} j\omega L_1 & -j\omega M \\ -j\omega M & j\omega L_2 \end{bmatrix} \tag{3.82}$$

Equating the matrix (3.82) to the impedance matrix of the unsymmetrical T-network of Fig. 3.9 we calculate its elements.

$$\begin{aligned} Z_2 &= j\omega M \\ Z_1 &= j\omega(L_1 - M) \\ Z_3 &= j\omega(L_2 - M) \end{aligned} \tag{3.83}$$

The result can be represented by the equivalent network of Fig. 3.31, which also gives the equivalent π-network of the transformer. The latter is found by taking the reciprocal of the matrix (3.82) and equating it to the admittance matrix of the π-network given in Fig. 3.4.

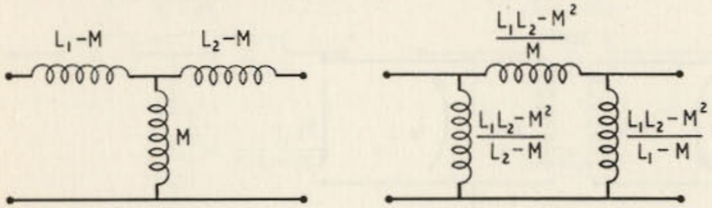


Fig. 3.31

The foregoing examples of equivalent networks suggest that there are limits beyond which it is not possible to make a configuration equivalent to a given circuit. Thus the transformer cannot be replaced by a symmetric T-network. The reason is understood when it is observed that the matrix (3.82) has unequal diagonal elements, hence it cannot be equal to a matrix with equal diagonal elements, which is characteristic of symmetric networks.

Having shown that certain circuit configurations can be equivalent to others, we now pass on to consider the second problem stated at the opening of this section. Given the parameters of a 'black box' is it possible to set up an actual circuit having the same parameters?

Before we show that it is, in fact, possible to do so, it may be appropriate to remark that in many practical problems it is convenient to think of a specific network configuration, rather than a vague 'black box'. Engineers often find it desirable to visualise an actual circuit rather than work in terms of abstract matrices, particularly in experimental situations where measurements are made or design problems solved.

At first let us introduce two typical examples: Limiting ourselves to passive networks, we assume a specific Z-matrix (symmetric) to be given as a statement of the characteristics of a 'black box'.

$$Z = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix}, \qquad z_{12} = z_{21}$$

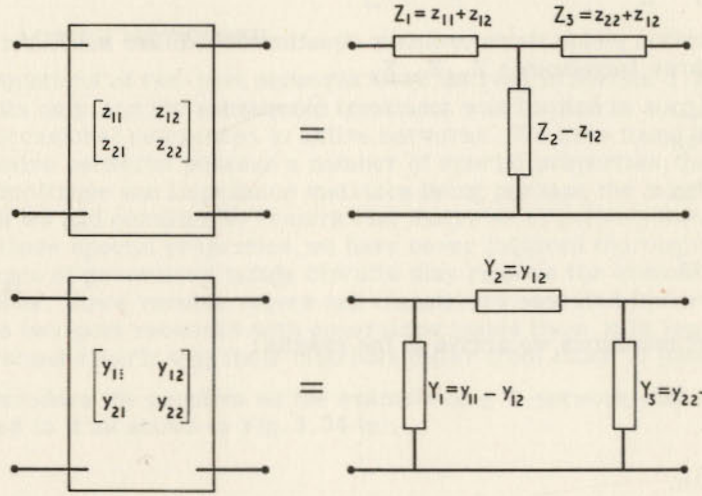


Fig. 3.32

We can easily construct a T-network having this matrix if we let

$$\begin{aligned} Z_2 &= -z_{12} = -z_{21} \\ Z_1 &= z_{11} - Z_2 = z_{11} + z_{12} \\ Z_3 &= z_{22} - Z_2 = z_{22} + z_{12} \end{aligned} \tag{3.84}$$

It is equally easy to set up a π-network to represent a given Y-matrix. Fig. 3.32 presents the solutions graphically.

The general procedure can be outlined as follows:

1. The parameters of a 'black box' are given.
2. The matrix of the same parameters for a desired equivalent circuit is written down.
3. The two matrices are equated, yielding equations which determine the equivalent circuit impedances in terms of the given parameters.

To illustrate the method we find the equivalent T-network of a 'black box'

whose h-parameters are known. The H-matrix of the T-network is written down starting from Fig. 3.9 and using Table 3.1.

$$\begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} = \begin{bmatrix} \dfrac{|Z|}{Z_2 + Z_3} & \dfrac{Z_2}{Z_2 + Z_3} \\ \dfrac{-Z_2}{Z_2 + Z_3} & \dfrac{1}{Z_2 + Z_3} \end{bmatrix} \tag{3.85}$$

This matrix equation yields three ordinary equations which are sufficient to determine the three impedances Z_1, Z_2, Z_3.

$$\begin{aligned} Z_2 + Z_3 &= \frac{1}{h_{22}} \\ \frac{Z_2}{Z_2 + Z_3} &= h_{12} \\ \frac{|Z|}{Z_2 + Z_3} &= h_{11} \end{aligned} \tag{3.86}$$

By successive eliminations we arrive at the results:

$$\begin{aligned} Z_2 &= \frac{h_{12}}{h_{22}} \\ Z_3 &= \frac{1 - h_{12}}{h_{22}} \\ Z_1 &= h_{11} + \frac{h_{12}\,(h_{12} - 1)}{h_{22}} \end{aligned} \tag{3.87}$$

which are summarised in Fig. 3.33. The reader should check that Equations 3.87 are dimensionally correct by referring back to Equation 3.48, and should satisfy himself that the H-matrix of the network of Fig. 3.33 is in fact identical with Equation 3.85.

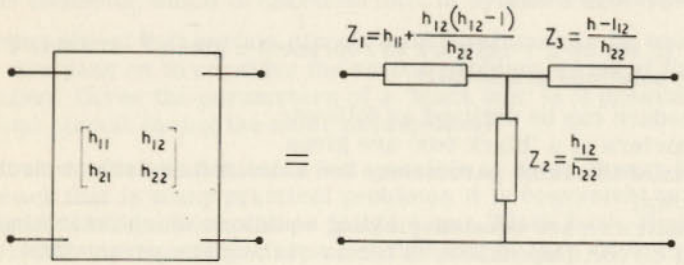


Fig. 3.33

As remarked before the object of having equivalent networks is to ease practical problems of circuit analysis, experimental or theoretical, and of circuit design. Once an equivalent circuit has been set up a generator can be applied to it and a load can be connected to its output terminals. This done, various electrical quantities can be measured, such as voltage and current ratios or input and output impedances. In this way an equivalent circuit may help to visualise a practical electrical problem, although a solution can always be effected on the basis of known parameters.

3.11 ACTIVE NETWORKS

The equations of two-port networks were derived in Section 3.2 for passive circuits only, and the subsequent treatment was limited to such circuits, with only occasional references to active networks. We have found that matrices of passive networks possess a number of special properties, the symmetry of the admittance and impedance matrices being perhaps the most notable. Although we had occasion to remark that matrices of active networks do not have these special properties, we have never inquired thoroughly how the presence of generators inside circuits may remove the symmetry of their matrices. Since vacuum valves and transistors operated linearly are equivalent to two-port networks with generators inside them, it is important to understand clearly why their matrices differ from those of passive networks.

We introduce the problem on the example on a T-network with an e.m.f. E applied to it as shown in Fig. 3.34 (a).

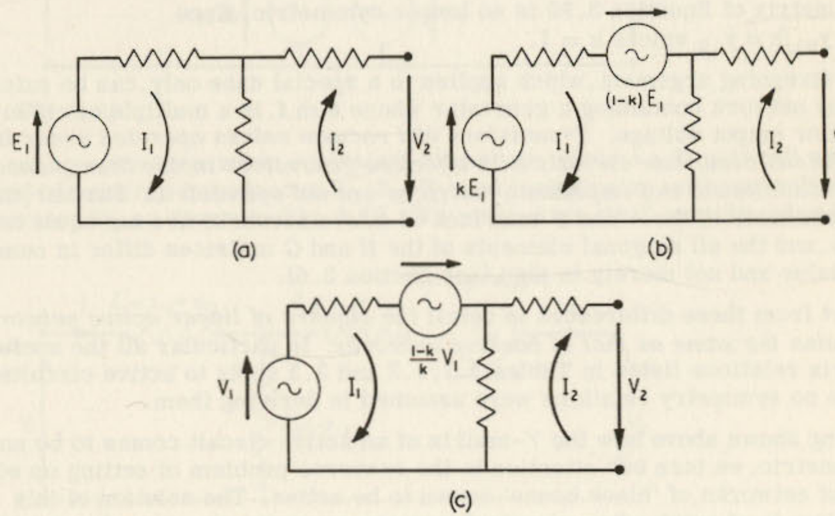


Fig. 3.34

The equations of this network can be written

$$I_1 = y_{11}\,E_1 + y_{12}\,V_2$$
$$I_2 = y_{21}\,E_1 + y_{22}\,V_2 \qquad (3.88)$$

where Y is the *symmetric* admittance matrix of the *passive* T-network of Fig. 3.34 (a). Since E_1 is effectively the vector sum of all generators in loop 1, we can assume that it consists of two fractions, kE_1 and $(1-k)\,E_1$, the latter placed inside the circuit, as shown in Fig. 3.34 (b). Equations 3.88 can now be rewritten in the form

$$\begin{aligned} I_1 &= y_{11}\Big[kE_1 + (1-k)E_1\Big] + y_{12}V_2 \\ I_2 &= y_{21}\Big[kE_1 + (1-k)E_1\Big] + y_{22}V_2 \end{aligned} \qquad (3.89)$$

Putting $kE_1 = V_1$ and taking it outside the square brackets we obtain the equations of the network of Fig. 3.34 (c).

$$\begin{aligned} I_1 &= \frac{y_{11}}{k}V_1 + y_{12}\,V_2 \\ I_2 &= \frac{y_{21}}{k}V_1 + y_{22}\,V_2 \end{aligned} \;,\quad \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} \dfrac{y_{11}}{k} & y_{12} \\ \dfrac{y_{21}}{k} & y_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \qquad (3.90)$$

The matrix of Equation 3.90 is no longer symmetric, since
$y_{21}/k \neq y_{12}$ unless $k = 1$.

The foregoing argument, which applies to a special case only, can be extended to any network containing a generator whose e.m.f. is a multiple of either the input or output voltage. *Transistors and vacuum valves operated under linear conditions constitute circuits with affective generators inside them, hence their admittance and impedance matrices are not symmetric.* Further, the determinants of the A and B matrices of active networks are not equal to unity, and the off diagonal elements of the H and G matrices differ in numerical value and not merely in sign (see Section 3.6).

Apart from these differences in detail *the algebra of linear active networks remains the same as that of passive networks*. In particular all the useful matrix relations listed in Tables 3.1, 3.2 and 3.3 apply to active circuits, since no symmetry relations were assumed in deriving them.

Having shown above how the Y-matrix of an active circuit comes to be non-symmetric, we turn our attention to the converse problem of setting up equivalent networks of 'black boxes' known to be active. The solution of this problem is of outstanding practical importance as it helps to understand the equivalent circuits of valves and transistors.

Two basic types of equivalent circuit can be used to represent an active two-port network.
1. Two-generator equivalent circuit.
2. One-generator equivalent circuit.

The fictitious generators may be either voltage or current sources.

We start with the equations using z-parameters which we rewrite here for convenience.

$$V_1 = z_{11}\,I_1 + z_{12}\,I_2, \; \left(z_{12} \neq z_{21}\right)$$
$$V_2 = z_{21}\,I_1 + z_{22}\,I_2 \qquad (3.91)$$

The first expression may be interpreted as equating the terminal voltage V_1 to a voltage drop $z_{11}\,I_1$ across an impedance z_{11}, plus a generator of voltage $z_{12}\,I_2$ opposing V_1. Similarly the second expression equates the terminal voltage V_2 to a drop $z_{22}\,I_2$ plus a voltage generator $z_{21}\,I_1$. Hence Equations 3.91 are satisfied by the circuit of Fig. 3.35, which is a two-generator equivalent circuit of the two-port network having the Z-matrix of Equations 3.91.

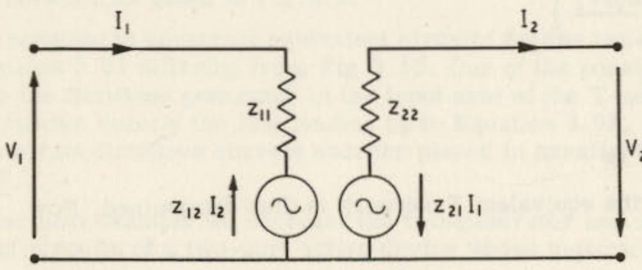


Fig. 3.35

To set up a one-generator equivalent circuit we assume a T-network composed of unknown impedances Z_1, Z_2, Z_3 and including an unknown voltage generator V_T as shown in Fig. 3.36. We write down its Z-matrix using the

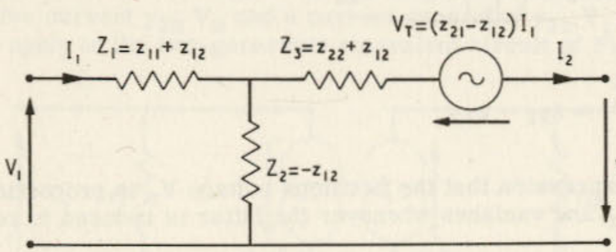


Fig. 3.36

open circuiting procedure, and equate it to the known matrix of Equations 3.91. We start with the transfer impedance z_{12}.

$$z_{12} = \left(\frac{V_1}{I_2}\right)_{I_1=0}$$

$$= \left(\frac{-Z_2 I_2}{I_2}\right)_{I_1=0}$$

$$= -Z_2$$

$$Z_2 = -z_{12}$$

This fixes the vertical branch of the T-network. Further

$$z_{11} = \left(\frac{V_1}{I_1}\right)_{I_2=0}$$

$$= \left(\frac{(Z_1 + Z_2)V_1}{V_1}\right)_{I_2=0}$$

$$= Z_1 + Z_2$$

whence $Z_1 = z_{11} + z_{12}$

The input branch of the equivalent T-network is thus determined. Now

$$z_{21} = \left(\frac{V_2}{I_1}\right)_{I_2=0} = \left(\frac{V_T - I_1 Z_2}{I_1}\right)_{I_2=0}$$

$$= \left(\frac{V_T}{I_1}\right)_{I_2=0} - Z_2 = \left(\frac{V_T}{I_1}\right)_{I_2=0} + z_{12}$$

and $$\left(\frac{V_T}{I_1}\right)_{I_2=0} = z_{21} - z_{12}$$

It follows from this expression that the fictitious voltage V_T is proportional to the input current I_1, and vanishes whenever the latter is reduced to zero. Hence we find

$$V_T = \left(z_{21} - z_{12}\right)I_1$$

Now it only remains to work out the impedance Z_3 from the equation

$$z_{22} = \left(\frac{V_2}{I_2}\right)_{I_1=0} = \left(\frac{(Z_2 + Z_3)I_2 + V_T}{I_2}\right)_{I_1=0} = Z_2 + Z_3 + \left(\frac{V_T}{I_2}\right)_{I_1=0}$$

$$= Z_2 + Z_3 + \left(\frac{(z_{21} - z_{12})I_1}{I_2}\right)_{I_1=0} = Z_2 + Z_3.$$

Hence $Z_3 = z_{22} + z_{12}$

The results may be summarised in the single matrix equation

$$\begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} = \begin{bmatrix} Z_1 + Z_2 & -Z_2 \\ \frac{V_T}{I_1} - Z_2 & Z_2 + Z_3 \end{bmatrix} \qquad (3.92)$$

where it must be remembered that V_T/I_1 is constant. It is instructive to compare the foregoing matrix of an active T-network with the matrix of a passive network, as given in Fig. 3.9.

It is possible to construct equivalent circuits for the two-port network of Equations 3.91 differing from Fig. 3.36. One of the possible solutions is to place the fictitious generator in the input arm of the T-network. The calculation follows exactly the line leading up to Equation 3.92. Other possibilities incorporate fictitious current sources placed in parallel with the branches of the T.

As our next example we consider the two-generator and one-generator equivalent circuits of a two-port active device whose y-parameters are known.

$$I_1 = y_{11}V_1 + y_{12}V_2, \quad \left(y_{12} \neq y_{21}\right)$$
$$I_2 = y_{21}V_1 + y_{22}V_2 \qquad (3.93)$$

I_1 can be regarded as the vector sum of a current $y_{11}\ V_1$, pushed through a passive admittance y_{11} by the voltage V_1, plus the current $y_{12}\ V_2$, generated by a constant current source of this magnitude. Similarly I_2 is the sum of a passive current $y_{22}\ V_2$ and a current generator $y_{21}\ V_1$. Hence Equations 3.93 apply to the two-generator equivalent circuit of Fig. 3.37.

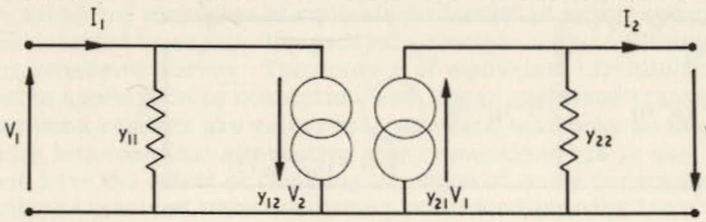


Fig. 3.37

The device of Equations 3.93 can be represented by a one-generator circuit in the form of a π-network, including a current source across its output terminals, as shown in Fig. 3.38. To evaluate the branch admittances and the current generator we follow the method that led up to Equation 3.92. Starting with the known transfer admittance y_{12} we find

$$y_{12} = \left(\frac{I_1}{V_2}\right)_{V_1=0}$$

$$= \left(\frac{Y_2 V_2}{V_2}\right)_{V_1=0}$$

$$= Y_2$$

The top branch of the π-network is thus determined. To find the input branch of the equivalent network we utilise the parameter y_{11}.

$$y_{11} = \left(\frac{I_1}{V_1}\right)_{V_2=0}$$

$$= \left(\frac{(Y_1 + Y_2)V_1}{V_1}\right)_{V_2=0}$$

$$= Y_1 + Y_2$$

$$Y_1 = y_{11} - y_{12}$$

The parameter y_{21}, together with the foregoing results, yields the magnitude of the fictitious current source which we denote by I_π.

$$y_{21} = \left(\frac{I_2}{V_1}\right)_{V_2=0} = \left(\frac{Y_2 V_1 - I_\pi}{V_1}\right)_{V_2=0} = Y_2 - \left(\frac{I_\pi}{V_2}\right)_{V_1=0}$$

$$\left(\frac{I_\pi}{V_1}\right)_{V_2=0} = y_{12} - y_{21}$$

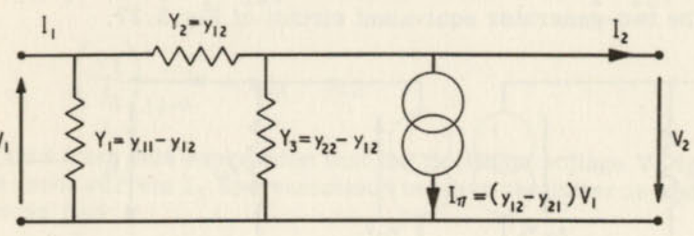


Fig. 3.38

This relation shows that the intensity of the fictitious current source is proportional to the input voltage V_1 and, therefore, vanishes when the latter is zero. Hence we find

$$I_\pi = \left(y_{12} - y_{21}\right)V_1$$

Finally the remaining shunt element of the equivalent circuit is found.

$$y_{22} = \left(\frac{I_2}{V_2}\right)_{V_1=0}$$

$$= \left(\frac{(Y_2 + Y_3)V_2 - I_\pi}{V_2}\right)_{V_1=0}$$

$$= Y_2 + Y_3 - \left(\frac{I_\pi}{V_2}\right)_{V_1=0}$$

Under conditions of short circuit at the input the last term of this expression vanishes. Hence

$$Y_3 = y_{22} - y_{12}$$

The method by which the single generator equivalent circuits have been derived is not unique. Alternative, more concise arguments can be used. For example, Equations 3.93 may be rearranged as follows:

$$\begin{aligned} I_1 &= y_{11} V_1 + y_{12} V_2 \\ I_2 &= y_{12} V_1 + y_{22} V_2 - \left(y_{12} - y_{21}\right)V_1 \end{aligned} \tag{3.94}$$

If, for the moment, we neglect the last term of the second equation, Equations 3.94 would represent a passive network, because their matrix would be symmetric. Hence they would have the equivalent π-network of Fig. 3.32 which is the same as the passive part of Fig. 3.38. The last term may be conceived to be a current source added across the output terminals of the passive circuit in opposition to the assumed direction of I_2, because of its negative sign. A similar argument, based on a rearrangement of equations, may be adduced to obtain the equivalent T-circuit of Fig. 3.36.

The foregoing examples of equivalent circuits of active two-port devices have been selected to explain the general principle, without attempting to provide a comprehensive survey. The concept of equivalent circuits finds its most extensive application in connection with linear valve and transistor circuits. Interested readers are referred to standard textbooks on the subject remembering however, that alternative sign conventions are in use. In general these have the effect of changing the signs of some parameters, so that care must be exercised when comparing results originating from different sources.

Part 2

Chapter 4

Developments of Matrix Algebra

The algebraic work of Chapter 2 will now be continued. In the present chapter the stock of rules and definitions will be extended, and problems that were beyond the power of elementary methods will be tackled. The subject matter to be covered will be adequate to deal with most electrical problems susceptible to a formulation in terms of matrices.

In the exposition of the more complicated subjects of rank, eigenvalue problems, functions of matrices, no attempt is made at mathematical rigour. Instead the treatment concentrates on the introduction of basic concepts, and a clear explanation of algebraic manipulations.

4.1 PARTITIONING OF MATRICES.

The elementary rules of matrix algebra explained in Chapter 2 lend themselves to a useful extension. Instead of handling individual matrix elements when products or sums are formed, it is possible to manipulate whole groups of elements. The matrices to be, say, multiplied together are first *partitioned*, and the resulting *submatrices* are then found to obey the same rules as individual elements.

We introduce the subject on the example of two matrices A and B, which we subdivide according to the following scheme.

$$A = \left[\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ \hline a_{31} & a_{32} & a_{33} \end{array}\right] = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$$

$$B = \left[\begin{array}{cc|c} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{array}\right] = \begin{bmatrix} B_1 & B_2 \end{bmatrix} \qquad (4.1)$$

The product AB is next worked out in terms of the submatrices shown in Equations 4.1 rather than individual elements.

$$AB = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \begin{bmatrix} B_1 & B_2 \end{bmatrix} = \begin{bmatrix} A_1B_1 & A_1B_2 \\ A_2B_1 & A_2B_2 \end{bmatrix} \qquad (4.2)$$

Forming the products A_iB_j in full we obtain the expected product matrix:

$$\left[\begin{array}{cc|c} a_{11}b_{11}+a_{12}b_{21}+a_{13}b_{31} & a_{11}b_{12}+a_{12}b_{22}+a_{13}b_{32} & a_{11}b_{13}+a_{12}b_{23}+a_{13}b_{33} \\ a_{21}b_{11}+a_{22}b_{21}+a_{23}b_{31} & a_{21}b_{12}+a_{22}b_{22}+a_{23}b_{32} & a_{21}b_{13}+a_{22}b_{23}+a_{23}b_{33} \\ \hline a_{31}b_{11}+a_{32}b_{21}+a_{33}b_{31} & a_{31}b_{12}+a_{32}b_{22}+a_{33}b_{32} & a_{31}b_{13}+a_{32}b_{23}+a_{33}b_{33} \end{array}\right] \qquad (4.3)$$

The dotted lines mark the extent of the subproducts A_iB_j.

Let us now attempt to form the reverse product of A and B retaining the partitions used above.

$$BA = \begin{bmatrix} B_1 & B_2 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \qquad = \begin{bmatrix} B_1A_1 + B_2A_2 \end{bmatrix}$$

$$\left[\begin{bmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} & b_{11}a_{13} + b_{12}a_{23} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} & b_{21}a_{13} + b_{22}a_{23} \\ b_{31}a_{11} + b_{32}a_{21} & b_{31}a_{12} + b_{32}a_{22} & b_{31}a_{13} + b_{32}a_{23} \end{bmatrix} + \begin{bmatrix} b_{13}a_{31} & b_{13}a_{32} & b_{13}a_{33} \\ b_{23}a_{31} & b_{23}a_{32} & b_{23}a_{33} \\ b_{33}a_{31} & b_{33}a_{32} & b_{33}a_{33} \end{bmatrix} \right] \qquad (4.4)$$

The dotted lines show the distinct submatrices which are to be added. After this has been done it is easily seen that the resulting matrix is of order 3×3 and is identical with the result of multiplying B by A in terms of their individual elements.

We conclude on the basis of these observations that the matrices A and B can be partitioned according to the scheme (Equations 4.1) without violating any of the basic rules of matrix algebra. It should also be noted that in general both the prefactor and the postfactor can be partitioned by either rows or columns. Thus, in Equations 4.2 the *rows* of the prefactor are partitioned but in Equations 4.4 the *columns* of the prefactor are partitioned.

Before proceeding to further extensions of the method of submatrices it is instructive to consider an example in which partitioning has been done incorrectly. Thus, let us write

$$P = \left[\begin{array}{c|c} p_{11} & p_{12} \\ p_{21} & p_{22} \\ p_{31} & p_{32} \end{array}\right] = \begin{bmatrix} P_1 & P_2 \end{bmatrix}$$

$$Q = \left[\begin{array}{cc|c} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \end{array}\right] = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \qquad (4.5)$$

Although the product PQ can be formed in terms of individual elements, it cannot be worked out in terms of the submatrices shown, since two row vectors cannot be multiplied together. Moreover the subproducts P_iQ_j cannot be formed either, because the orders of the submatrices are not properly matched.

The above example is indicative of a rule concerning partitioning. If one factor matrix is partitioned according to rows, the other must be partitioned according to columns, and vice versa. It should be noted that the matrices A and B in Equations 4.1 conform to this rule. The rule is further illustrated by Equations 4.5 if we partition the matrix Q by rows. The product

$$\begin{bmatrix} P_1 & P_2 \end{bmatrix} \begin{bmatrix} Q_3 \\ Q_4 \end{bmatrix} = \begin{bmatrix} P_1 Q_3 + P_2 Q_4 \end{bmatrix} \tag{4.6}$$

can then be worked out, yielding the correct result for PQ.

The following is another example of partitioning which does not work:

$$A = \left[\begin{array}{c:cc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{array}\right], \quad B = \left[\begin{array}{ccc} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ \hdashline b_{31} & b_{32} & b_{33} \end{array}\right] \tag{4.7}$$

The product AB cannot be formed according to the partitioning, despite the fact that one matrix is divided by rows and the other by columns. It is said that the partitions are inserted *unconformably*, or in other words the orders of the submatrices are not properly matched. On the other hand, the reverse product BA is quite feasible as the reader can verify.

The last example is typical of the general rule that the prefactor can be partitioned by rows in any manner, and the postfactor can be partitioned by columns in any manner, always yielding submatrices which can be multiplied according to the rules of matrix algebra.

So far we have considered cases where a given matrix has been partitioned either by rows or by columns, but not both simultaneously. To see that the latter is possible, we consider two matrices P and Q, not necessarily square, which we intend to multiply PQ. We partition the first matrix by rows and the second by columns in any way whatever, as is possible by the preceding rule. Next we superimpose partitions on the columns of P and the rows of Q which are conformable. The resulting submatrices of P and Q are conformable, and the partitioned multiplication can be carried out.

$$PQ = \left[\begin{array}{cc:c} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ \hdashline p_{31} & p_{32} & p_{33} \end{array}\right] \left[\begin{array}{cc:cc} q_{11} & q_{12} & q_{13} & q_{14} \\ q_{21} & q_{22} & q_{23} & q_{24} \\ \hdashline q_{31} & q_{32} & q_{33} & q_{34} \end{array}\right] = \begin{bmatrix} P_1 & P_3 \\ P_3 & P_4 \end{bmatrix} \begin{bmatrix} Q_1 & Q_2 \\ Q_3 & Q_4 \end{bmatrix} \tag{4.8}$$

The reader should work out all subproducts in Equation 4.8 and check that the results agree with element by element multiplication of P by Q.

A justification for the partitioning procedure is to be sought in the general summations (see Section 2.4) that appear when products of matrices are formed. Thus the typical element of the product

$$R = PQ \tag{4.9}$$

is given by

$$r_{ij} = \sum_{k=1}^{n} p_{ik} q_{kj} \tag{4.10}$$

where n is the number of columns in P and rows in Q. The sum, Equation 4.10, can be split into a number of partial sums, each extending over a fraction of the range from 1 to n.

$$r_{ij} = \sum_{k=1}^{k_1} p_{ik} q_{kj} + \sum_{k=k_1+1}^{k_2} p_{ik} q_{kj} + \cdots + \sum_{k=k_1+1}^{n} p_{ik} q_{kj} \tag{4.11}$$

Now, every partial sum in Equation 4.11 can be considered to stand for the typical element of a product of submatrices taken out of P and Q. What interests us here is the method of selecting these submatrices. The subscript k labels the columns of the prefactor and the rows of the postfactor, therefore each partial sum contains elements from columns of P and rows of Q having the same labels. This is the conformable partitioning of P and Q by columns and rows respectively, as illustrated in Equation 4.8. Moreover the individual elements of the product matrix $\left[r_{ij}\right]$ are given by separate sums of the form of Equation 4.10, hence they can be computed in any sequence of the suffices i and j. This reflects the arbitrary partitioning of the prefactor by rows and the postfactor by columns, once again exemplified by Equation 4.8.

In connection with the Equation 4.11 it should be observed that the subranges of k need not be consecutive at all. In view of this we conclude that the concept of submatrices is, in fact, more general than disclosed by the examples considered. Submatrices can be formed from rows and columns which are scattered over the original matrix, as well as from those which adjoin each other.

It is hoped that the introduction of partitioned matrices in the present section has provided the reader with some more experience of manipulating matrices, but this is not by any means the sole object of the excercise. Partitioning is a further step in the development of the concise algebra of matrices and we shall soon see how it helps to reduce complicated general expressions to manageable form.

4.2 ELEMENTARY OPERATIONS ON ROWS AND COLUMNS OF A MATRIX

The manipulations that come under this heading include multiplication of a row or column by a constant factor, interchange of rows or columns within a matrix, addition of one row or column to another etc... The object of the present section is to show how to carry out such operations using the elementary rules of matrix algebra laid down in Chapter 2.

We start with the problem of how to multiply a single row of a matrix P by a constant factor, without affecting the remaining rows. To effect this operation we try the following product:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ kp_{21} & kp_{22} & kp_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} \qquad (4.12)$$

Hence, to multiply a row of a matrix by a constant factor we premultiply the matrix by the unit matrix with the corresponding diagonal element replaced by k. The rule is readily extended to include simultaneous multiplication of several rows of a matrix by a succession of scalars k_i.
Example:

$$\begin{bmatrix} k_1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & k_3 & 0 \\ 0 & 0 & 0 & k_4 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \\ p_{31} & p_{32} \\ p_{41} & p_{42} \end{bmatrix} = \begin{bmatrix} k_1p_{11} & k_1p_{12} \\ p_{21} & p_{22} \\ k_3p_{31} & k_3p_{32} \\ k_4p_{41} & k_4p_{42} \end{bmatrix} \qquad (4.13)$$

To multiply columns of a matrix by a constant we use the appropriate diagonal matrix in postmultiplication, as shown below.

$$\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \\ p_{31} & p_{32} \end{bmatrix} \begin{bmatrix} k_1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} k_1p_{11} & p_{12} \\ k_1p_{21} & p_{22} \\ k_1p_{31} & p_{32} \end{bmatrix} \qquad (4.14)$$

To *interchange two rows* of a matrix we premultiply it by the unit matrix with the *same rows interchanged*. Example:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ c_1 & c_2 & c_3 & c_4 \\ b_1 & b_2 & b_3 & b_4 \end{bmatrix} \qquad (4.15)$$

The reader should go over this multiplication carefully, to make sure that the mechanism of interchange is clearly seen.

To *interchange columns* of a matrix, a unit matrix with *interchanged columns* is used in postmultiplication.

$$\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_4 & b_4 & c_4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} a_1 & c_1 & b_1 \\ a_2 & c_2 & b_2 \\ a_3 & c_3 & b_3 \\ a_4 & c_4 & b_4 \end{bmatrix} \qquad (4.16)$$

Here again the mechanism of interchange should be carefully studied by carrying out the multiplication in full.

Next consider the product

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} = \begin{bmatrix} a_1 + c_1 & a_2 + c_2 & a_3 + c_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \qquad (4.17)$$

It will be observed that the nett result of the above multiplication is to add the third row to the first row of the postfactor. The operative element in this manipulation is the extra unit in the top right hand corner of the prefactor. For purposes of memorising, it is useful to note that the extra element is in the first row, but directly above the unit in the third row, of the prefactor. Symbolically this can be taken to mean that the third row is to be added to the first.

Simultaneous addition of several rows of a matrix to a given row can be accomplished by inserting units into the appropriate positions of the unit matrix. Thus the matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

when used in premultiplication, will cause the addition of the first and second rows to the third row of the postfactor.

Example:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_2 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \\ a_1 + b_1 + c_1 & a_1 + b_2 + c_2 \end{bmatrix} \qquad (4.18)$$

To effect the addition of columns of a matrix rather than rows, suitably modified unit matrices are used in postmultiplication. Example:

$$\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} a_1 & a_1 + b_1 + c_1 & c_1 \\ a_2 & a_2 + b_2 + c_2 & c_2 \end{bmatrix} \tag{4.19}$$

Sometimes it is required to add a multiple of a column or row to another. This is done by matrices similar to the above with the difference that the appropriate multiplier is placed in an off diagonal position, instead of a unit. Thus the matrix

$$\begin{bmatrix} 1 & k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

when applied in premultiplication will add k×row 2 to row 1, and when used in postmultiplication it will add k×column 1 to column 2, as the reader may verify on an example.

It should be observed that although all the above rules of elementary operations have been stated for both rows and columns, the latter can always be deduced from the former by transposition, and vice versa. Thus we have the example

$$\begin{bmatrix} 1 & k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_2 \end{bmatrix} = \begin{bmatrix} a_1 + kb_1 & a_2 + kb_2 \\ b_1 & b_2 \\ c_1 & c_2 \end{bmatrix} \tag{4.20}$$

and its transpose

$$\begin{bmatrix} a_1 & b_1 & c_1 \\ b_1 & b_2 & c_2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ k & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_1 + kb_1 & b_1 & c_1 \\ a_2 + kb_2 & b_2 & c_2 \end{bmatrix} \tag{4.21}$$

From the distinct elementary operations described so far, composite ones can be formed by multiplying the appropriate matrices together. The product

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & k & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.22}$$

when used in premultiplication, will multiply the second row of a desired matrix by k and add it to the first row. The reader should check this on an example.

At this stage an important observation should be made regarding the determinants of square matrices, on which elementary operations have been performed. A reference back to Section 2.13 will disclose that the properties listed there are all concerned with determinants of matrices subjected to the operations discussed above. It should be particularly noted for future reference that such operations can affect the value of a determinant only by a constant factor. *Thus, a non-zero determinant cannot be made to vanish by any number of elementary operations on its rows or columns. Conversely, a vanishing determinant cannot be made to assume a finite value by such operations.*

A further important fact should be noted. The matrices effecting elementary operations are non-singular, hence their reciprocals can always be formed. Moreover, *the reciprocal matrix effects the reverse operation.* The reader should go over the examples given above to satisfy himself that this is so.

The rules concerning elementary operations, outlined in this section, are easily extended to partitioned matrices. Since the latter behave under multiplication in exactly the same way as non-partitioned matrices, all we have to do is to make sure that the partitioning is conformable. As an example we take the 4 × 2 matrix

$$\begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_2 \\ d_1 & d_2 \end{bmatrix} \tag{4.23}$$

and set ourselves the problem of interchanging the first two rows, taken bodily together, with the third. The fourth row is to remain unchanged. The student will appreciate that this can be accomplished by means of two operations on individual rows, successively shifting the c_i row upwards. However, we can obtain the same result in one step, if we partition the matrix 4.23 suitably.

Let us write

$$\begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \\ \hdashline c_1 & c_2 \\ \hdashline d_1 & d_2 \end{bmatrix} = \begin{bmatrix} A \\ C \\ D \end{bmatrix} \tag{4.24}$$

and

$$\left[\begin{array}{cc|c|c} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] = \begin{bmatrix} I_1 & I_2 & I_3 \end{bmatrix} \tag{4.25}$$

Partitioned premultiplication of Equation 4.24 by Equation 4.25 gives

$$\begin{bmatrix} I_1 & I_2 & I_3 \end{bmatrix} \begin{bmatrix} A \\ C \\ D \end{bmatrix} = \begin{bmatrix} I_1A + I_2C + I_3D \end{bmatrix}$$

$$= \left[\begin{bmatrix} 0 & 0 \\ a_1 & a_2 \\ b_1 & b_2 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} c_1 & c_2 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ d_1 & d_2 \end{bmatrix} \right] = \begin{bmatrix} c_1 & c_2 \\ a_1 & a_2 \\ b_1 & b_2 \\ d_1 & d_2 \end{bmatrix} \tag{4.26}$$

The importance of the elementary operations described in the foregoing section lies in the fact that they can be used to reduce a matrix to equivalent diagonal form. This problem forms the subject of the following section.

4.3 EQUIVALENT MATRICES

In many important problems of applied mathematics it is necessary to transform or reduce a given matrix to diagonal form. Problems of this type come up for discussion in Chapter 5, in connection with applications. In the present chapter the requisite algebraic methods are explained, and in this section it is shown how a matrix can be reduced to diagonal form by a succession of elementary operations on its rows and columns.

The method is best introduced on a numerical example, say the following matrix of order 3×3

$$\begin{bmatrix} 1 & 3 & 2 \\ 4 & 6 & 1 \\ 5 & 7 & 0 \end{bmatrix} \tag{4.27}$$

As a first step we add ($-\,^5/_4 \times$ row 2) to row 3, which is effected by the product

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -^5/_4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 2 \\ 4 & 6 & 1 \\ 5 & 7 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 2 \\ 4 & 6 & 1 \\ 0 & -^1/_2 & -^5/_4 \end{bmatrix}$$

Next we subject the result to the operation (row 2 $- 4 \times$ row 1).

$$\begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 2 \\ 4 & 6 & 1 \\ 0 & -^1/_2 & -^5/_4 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -6 & -7 \\ 0 & -^1/_2 & -^5/_4 \end{bmatrix}$$

As the third step we add (row 3 $- \,^1/_{12} \times 2$).

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -^1/_{12} & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 2 \\ 0 & -6 & -7 \\ 0 & -^1/_2 & -^5/_4 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -6 & -7 \\ 0 & 0 & -^2/_3 \end{bmatrix} \tag{4.28}$$

We have thus reduced to zero the elements lying below the principle diagonal of the original matrix. This has been accomplished in steps of elementary operations with the overall effect given by the product

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -^1/_{12} & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -^5/_4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ ^1/_3 & -^4/_3 & 1 \end{bmatrix} \tag{4.29}$$

The reader should satisfy himself that a direct multiplication of matrix (4.27) by matrix (4.29) will yield the matrix on the right hand side of Equation 4.28.

The mechanism of reducing the off diagonal elements to zero should be carefully studied. At first the element in the bottom left hand corner is made equal to zero by adding to the third row a suitable multiple of the second row. Then the element directly above is reduced to zero by adding to the second row a selected multiple of the first row. In this way all elements of the first column below the principal diagonal are reduced to zero. The bottom element of the second column is made to vanish by the addition of the right multiple of the second row to the third. This final step naturally leaves the zeros in the first column unaffected.

The foregoing method of reducing off diagonal elements of a matrix to zero is not limited to our example of order 3×3, but is quite general.

The elements above the principal diagonal of our matrix can be made to vanish by either of two methods. First, we can go on operating on the rows by premultiplication with suitable matrices. Some work on the above example will satisfy the reader that this is possible.

Alternatively we can operate on the columns of the matrix on the right hand side of Equation 4.28 by multiplying with selected matrices. Thus, the top right hand corner element is removed by the operation (column 3 $-\,^2/_3 \times$ column 2).

$$\begin{bmatrix} 1 & 3 & 2 \\ 0 & -6 & -7 \\ 0 & 0 & -2/3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -2/3 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 0 \\ 0 & -6 & -3 \\ 0 & 0 & -2/3 \end{bmatrix}$$

In the next stage the second element of the first row is removed by the operation (column 2 − 3 × column 1).

$$\begin{bmatrix} 1 & 3 & 0 \\ 0 & -6 & -3 \\ 0 & 0 & -2/3 \end{bmatrix} \begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -6 & -3 \\ 0 & 0 & -2/3 \end{bmatrix}$$

Finally, the second element in the third column is removed by the operation (column 3 − ½ × column 2).

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -6 & -3 \\ 0 & 0 & -2/3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1/2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -6 & 0 \\ 0 & 0 & -2/3 \end{bmatrix} \tag{4.30}$$

We have thus reduced the original matrix (4.27) to the diagonal form, Equation 4.30, by means of elementary operations in the shape of pre- and postmultiplication by suitable matrices. The resulting matrix used in postmultiplication is

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -2/3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1/2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -3 & 3/2 \\ 0 & 1 & -7/6 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.31}$$

Let us now summarise our results by writing down the relation transforming the original matrix (4.27) into the matrix on the right hand side of Equation 4.30, using the matrices of Equations 4.29 and 4.31.

$$\begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 1/3 & -4/3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 2 \\ 4 & 6 & 1 \\ 5 & 7 & 0 \end{bmatrix} \begin{bmatrix} 1 & -3 & 3/2 \\ 0 & 1 & -7/6 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -6 & 0 \\ 0 & 0 & -2/3 \end{bmatrix} \tag{4.32}$$

The conclusion to be drawn from this example is that a matrix A can be reduced to a diagonal form D by a transformation of the form

$$PAQ = D$$

The diagonal matrix D, obtained from A by means of a succession of elementary operations, is said to be equivalent to it.

More generally, two matrices A and B, related by the transformation

$$B = PAQ \tag{4.33}$$

are said to be *equivalent*. Equation 4.33 is referred to as an *equivalence transformation*.

As pointed out on p. 123 the transforming matrices P and Q are non-singular. Hence we can write

$$A = P^{-1}BQ^{-1}$$

where P^{-1} and Q^{-1} reverse the elementary operations originally applied to A.

We can go a step further, beyond the result of Equation 4.32 and reduce our matrix to a unit matrix by a suitable pre- or postmultiplication. Proceeding by premultiplication we find:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1/6 & 0 \\ 0 & 0 & -3/2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & -6 & 0 \\ 0 & 0 & -2/3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.34}$$

Since diagonal matrices commute under multiplication (see Section 2.11), the same matrix as in Equation 4.34 can be used in postmultiplication. Hence the final stage of reduction can be absorbed into either P or Q, of Equation 4.33.

It should be noted in the foregoing example that all the diagonal positions of the reduced forms, Equations 4.32 or 4.34 are occupied by non zero elements. This is not always the case, as the reader can readily verify by working through the following example.

$$\begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 1/3 & -4/3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 1 \\ 4 & 6 & -1 \\ 5 & 7 & -5/3 \end{bmatrix} \begin{bmatrix} 1 & -3 & 3/2 \\ 0 & 1 & -5/6 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -6 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{4.35}$$

This result should prevent the conclusion that all square matrices of the same order are equivalent, because they can be reduced to the unit matrix with the help of the method which led up to Equation 4.34. Equation 4.35 demonstrates that in certain cases some diagonal positions turn out to be vacant. However many elementary operations are used, the unit matrix of Equation 4.34 cannot be transformed into that of Equation 4.35, hence the two matrices are not equivalent.

The importance of these observations will become apparent in subsequent sections.

It is instructive to note at this point an important property of the determinants of equivalent square matrices. Since equivalent matrices are related by a succession of elementary operations, their determinants can only differ by a constant factor as was shown on page 123. In particular, if it is found that the determinant of a matrix vanishes, we can conclude that the determinants of all matrices equivalent to it also vanish. For example, a diagonal form may be known. It is then obvious whether its determinant vanishes or not, and the information is immediately applicable to all equivalent forms.

The method of reducing a matrix to equivalent form, outlined in this section, applies to rectangular matrices as well as square ones. If a matrix of order $m \times n$ is to be reduced, square matrices of order $m \times m$ must be used in premultiplication, and of order $n \times n$ in postmultiplication. Hence the reduced matrix will be of order $m \times n$, as the original one. Example of order $m = 2$ and $n = 3$:

$$\begin{bmatrix} 1 & 1 \\ -2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 3 & 0 \\ 2 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & -3 & -18 \\ 0 & 1 & 6 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} \tag{4.36}$$

Partitioned matrices are within the scope of the above methods with some reservations to be stated later. We illustrate the procedure on a *square non-singular* matrix.

Let us take

$$A = \begin{bmatrix} A_1 & A_2 \\ B_1 & B_2 \end{bmatrix} \tag{4.37}$$

where A_1 is square, and operate on rows as follows:

$$\begin{bmatrix} I & 0 \\ -B_1A_1^{-1} & I \end{bmatrix} \begin{bmatrix} A_1 & A_2 \\ B_1 & B_2 \end{bmatrix} = \begin{bmatrix} A_1 & A_2 \\ 0 & -B_1A_1^{-1}A_2 + B_2 \end{bmatrix} \tag{4.38}$$

A suitable postmultiplication removes the top right hand element of the right side of Equation 4.38 by operating on columns.

$$\begin{bmatrix} A_1 & A_2 \\ 0 & -B_1A_1^{-1}A_2 + B_2 \end{bmatrix} \begin{bmatrix} I & -A_1^{-1}A_2 \\ 0 & I \end{bmatrix} = \begin{bmatrix} A_1 & 0 \\ 0 & -B_1A_1^{-1}A_2 + B_2 \end{bmatrix} \tag{4.39}$$

In Equation 4.39 we have a matrix reduced to partitioned diagonal form, which in general will not be diagonal in terms of its individual elements. If it is desired to reduce A to diagonal form in terms of its elements, an additional operation may be applied to Equation 4.39, leaving as a result unit matrices in diagonal positions.

$$\begin{bmatrix} A_1^{-1} & 0 \\ 0 & (-B_1A_1^{-1}A_2 + B_2)^{-1} \end{bmatrix} \begin{bmatrix} A_1 & 0 \\ 0 & -B_1A_1^{-1}A_2 + B_2 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{4.40}$$

Here the prefactor can be absorbed into the prefactor of Equation 4.38, so that the whole reducing operation assumes the form

$$PAQ = I$$

4.4 THE RANK OF A MATRIX

In the preceding section we had occasion to point out that not all matrices of the same order are equivalent. In the present section we introduce the important concept of rank of a matrix, and show how it can be used to identify and classify equivalent matrices, a problem often encountered in applications of matrix algebra.

The rank of a matrix is a number equal to the order of the highest order non-vanishing minor, that can be formed from the matrix. The rank is usually denoted by the letter r. The definition applies to both square and rectangular matrices.

To illustrate the definition we consider some numerical examples.

Square matrix of order 3×3:

$$\begin{bmatrix} 2 & 0 & 6 \\ 4 & 7 & 12 \\ 1 & 5 & 3 \end{bmatrix}$$

The determinant of this matrix vanishes, as can be seen on subtracting ($3 \times$ column 1) from column 3. Hence its rank is less than its order - $r < 3$. Next we proceed to evaluate minors of order 2×2 to see whether they vanish. The top left hand minor equals 14, hence the rank of the matrix is $r = 2$.

Square matrices of order 2×2:

$$\begin{bmatrix} 1/2 & 2 \\ 3 & 12 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$$

The determinant of the first matrix vanishes, hence its rank is $r = 1$. The determinant of the second matrix equals -2, hence its rank is the same as its order, $r = 2$.

Rectangular matrices of order 3×4

$$\begin{bmatrix} 1 & 0 & 2 & 0 \\ 3 & 4 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix}, \quad \begin{bmatrix} 2 & 0 & 6 & 0 \\ 4 & 7 & 12 & 14 \\ 1 & 5 & 3 & 10 \end{bmatrix}$$

On evaluating the left hand minor of the first matrix it is seen that it equals 4, hence the rank is $r = 3$. All possible minors of order 3×3 taken out of the second matrix vanish, therefore its rank must be $r < 3$. Since the top left hand minor of order 2×2 does not vanish the rank is $r = 2$.

The above example demonstrates that the rank of a rectangular matrix can at most equal the number of its rows or columns, whichever is less.

The rank of diagonal matrices is easily ascertained by inspection. Thus the matrix

$$\begin{bmatrix} a & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & b \end{bmatrix}$$

is of rank $r = 2$, because the minor of order 2×2 obtained by striking out the second row and column is ab. It is assumed that both a and b are non zero numbers.

A useful fact regarding the rank of matrices subject to elementary operations can now be established.

We have noted on pages 123 and 128 two important properties of determinants of square matrices. First, the vanishing or otherwise of the determinant remains unaffected under elementary operations. Second, the fact that the determinant of a matrix vanishes or not applies to all equivalent matrices.

From the foregoing we can draw two limited conclusions. First, when the rank of a square matrix equals its order $r = n$, the fact is unaltered by elementary operations. Second, when the rank is less than the order of the matrix $r < n$, this fact is also unaltered by elementary operations. The same observations apply to equivalent matrices, since they are related by elementary operations.

It is possible to establish a much more far reaching and useful theorem concerning the rank of matrices under elementary operations. We shall now show that *the rank, whatever its value, of any matrix, square or rectangular, remains unchanged under elementary operations. In other words, equivalent matrices have the same rank.*

To prove the theorem we must consider what happens to the minors when the rows and columns of a matrix are subjected to elementary operations.

It is important to note that minors are affected more severely than determinants of square matrices by the operation of adding multiples of some rows or columns to a given row or column. Whereas for determinants such operations are covered by Rule 5 p. 49 this does not apply to minors. The operation

$$\text{row}_i + \lambda_1 \text{row}_1 + \lambda_2 \text{row}_2 + \cdots + \lambda_k \text{row}_k \tag{4.41}$$

may introduce into a minor containing the i-th row, the elements of a row it did not contain originally. Despite this difficulty we can show that such operations do not affect the vanishing of minors.

We consider a matrix A and single out for attention one of its minors containing the i-th row. The minor, which is assumed to be of order $(r + 1) \times (r + 1)$, will be denoted by the symbol $|A_i|$. The matrix A is now subjected to the operation (4.41). As a result the minor containing the i-th row of the *modified* matrix will be much more complicated than $|A_i|$. In fact this complicated minor, when expanded in terms of the elements of its i-th row, will have the value

$$|A_i| + \lambda_1 |A_1| + \lambda_2 |A_2| + \cdots + \lambda_k |A_k| \tag{4.42}$$

where $|A_1|, |A_2|$ etc. are minors which differ from $|A_i|$ by having the 1st, 2nd etc. row substituted for the i-th row. It follows from (4.42) that if all minors of order $(r + 1) \times (r + 1)$, taken out of the matrix A, vanish, then all minors of the same order, taken out of the matrix modified by the operation (4.41), also vanish. Similar remarks apply when columns are subjected to the operation (4.41).

The other elementary operations can affect only the sign of a minor but not its numerical value. Hence, if all minors of order $(r + 1) \times (r + 1)$ in A vanish, then all minors of the same order in $B = P\,A\,Q$ also vanish, where P and Q represent elementary operations. Therefore the rank of B cannot exceed the rank of A.

As it can be shown (see p. 127) that P^{-1} and Q^{-1} also represent elementary operations, the same conclusion applies to $A = P^{-1} B Q^{-1}$. Hence A and B must have the same rank.

The reader should retrace the steps of the foregoing argument on the example of a rectangular matrix of low order.

Using the concept of rank, and the above theorem on the rank of equivalent matrices, it is possible to classify matrices according to the form they assume after reduction to diagonal form.

As a first step we reduce to equivalent diagonal form an arbitrary matrix A of order $m \times n$ and of rank $r \leqslant m, n$, using the procedure for partitioned matrices outlined on p. 128. A being of rank r, the greatest non-vanishing

minor that can be found inside it is of order r × r. The reduction starts with interchanges of rows and columns of A, designed to bring such a nonvanishing minor into the top left hand corner. The rank of A remains, of course, unchanged under this operation. This accomplished, A is partitioned according to the scheme.

$$A = \begin{bmatrix} A_1 & A_2 \\ B_1 & B_2 \end{bmatrix} \tag{4.43}$$

where A_1 is of order r × r and therefore non-singular. Next we remove the submatrix B_1 by means of a suitable premultiplication.

$$\begin{bmatrix} I & 0 \\ -B_1A_1^{-1} & I \end{bmatrix}\begin{bmatrix} A_1 & A_2 \\ B_1 & B_2 \end{bmatrix} = \begin{bmatrix} A_1 & A_2 \\ 0 & -B_1A_1^{-1}A_2 + B_2 \end{bmatrix} \tag{4.44}$$

At this point we interrupt the reduction procedure to show that the submatrix in the bottom right hand corner of Equation 4.44 is a null matrix. For, if this were not the case, we could use a non-zero element from this submatrix and join it on to A_1, together with a row of zeros and a column from A_2. E.G. for A_1 of order r × r = 2 × 2 we would find

$$\begin{bmatrix} a_{11} & a_{12} & p \\ a_{21} & a_{22} & q \\ 0 & 0 & d \end{bmatrix}$$

The determinant of this submatrix of order (r + 1) × (r + 1) = 3 × 3 does not vanish, contrary to the hypothesis that the original matrix A was of rank r. Hence we must accept it that the submatrix $-B_1A_1^{-1}A_2 + B_2$ consists entirely of zeros, and therefore the right hand side of Equation 4.44 is really of the form

$$\begin{bmatrix} A_1 & A_2 \\ 0 & 0 \end{bmatrix} \tag{4.45}$$

With the help of a postmultiplication (4.45) can be reduced to diagonal form in terms of submatrices.

$$\begin{bmatrix} A_1 & A_2 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} I & -A_1^{-1}A_2 \\ 0 & I \end{bmatrix} = \begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix} \tag{4.46}$$

As a final step in the reduction of A to equivalent diagonal form, Equation 4.46, right hand side, can be made to have unit elements in the diagonal positions of A_1.

$$\begin{bmatrix} A_1^{-1} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \tag{4.47}$$

At each stage of the foregoing argument the reader should check that all the products of submatrices can actually be formed. Moreover, at no stage was it necessary to take reciprocals except A^{-1}.

The result (Equation 4.47) should be compared with the closing paragraphs of the preceding section. There it was stipulated that the matrix dealt with, was both square and nonsingular. This was equivalent to the assumption that its rank was equal to its order, and consequently the resulting diagonal form was a unit matrix. By contrast, Equation 4.47 applies to any matrix of rank r, square or rectangular.

From Equation 4.47 we see that every matrix of rank r can be reduced to an equivalent form having exactly r non-zero elements in its leading diagonal positions. Moreover, since all matrices of the same order and rank can be reduced to the same form, they are equivalent. It is perhaps in this connection that the rather abstract concept of rank assumes its most concrete shape. The reader is encouraged to visualise the form of Equation 4.47 whenever the rank of a matrix is mentioned.

4.5 SOLUTION OF HOMOGENEOUS LINEAR EQUATIONS

We are now sufficiently equipped with methods of matrix algebra to tackle the solution of systems of homogeneous linear equations. The reader's attention was drawn to this problem in Chapter 2 (see p. 58), where it was emphasised that the methods available then were inadequate to attempt a solution.

We start straight away with the general case in which the number of unknowns is not the same as the number of equations, so that the matrix of the system is rectangular. The problem will be considered on the example of a system for which the number of equations is m = 3 and the number of unknowns is n = 4. The matrix of the system is assumed to be of rank r = 2. In parallel with the example we shall write out the general case.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 &= 0 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 &= 0 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 &= 0 \\ A\,X &= O \end{aligned} \tag{4.48}$$

The equations are assumed to be so arranged that the minor of order 2 × 2 in the top left hand corner does not vanish. It is then possible to rewrite Equations 4.48 in the partitioned form

$$\left[\begin{array}{cc:cc} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hdashline a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right]\left[\begin{array}{c} x_1 \\ x_2 \\ \hdashline x_3 \\ x_4 \end{array}\right] = 0$$

$$\begin{bmatrix} A_1 & A_2 \\ B_1 & B_2 \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = 0 \qquad (4.49)$$

Both sides of Equation 4.49 can now be premultiplied by a matrix P, selected according to the principle explained at the close of the preceding section. As a result Equation 4.49 assumes the form

$$PAX = O$$

$$\begin{bmatrix} A_1 & A_2 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = 0 \qquad (4.50)$$

This means that the number of equations has been effectively reduced. Writing our example in full, we are left with only 2 equations.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 &= 0 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 &= 0 \end{aligned} \qquad (4.51)$$

The reader should note that the linear operation on rows of A, performed by the matrix P above, is equivalent to the familiar procedure of adding multiples of the first two equations to the third. In our example the latter vanishes.

Equations 4.51 are next rewritten in the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= -a_{13}x_3 - a_{14}x_4 \\ a_{21}x_1 + a_{22}x_2 &= -a_{23}x_3 - a_{24}x_4 \\ A_1X_1 &= -A_2X_2 \end{aligned} \qquad (4.52)$$

To solve Equations 4.52 we substitute any numbers for x_3 and x_4 and reduce the problem to the simple form dealt with in Chapter 2. The values of x_1 and x_2 are

$$X_1 = -A_1^{-1} A_2X_2 \qquad (4.53)$$

The reciprocal A_1^{-1} can be formed because we assumed that the rank of the system was 2, and a non-vanishing minor of order 2×2 was brought into the leading position.

The values of the unknowns thus found clearly satisfy Equation 4.52, but to form a solution of the original system (Equation 4.48), they must also satisfy the last equation. This we now verify in the general case by direct substitution. Referring back to Equation 4.49 we must show that the equation

$$B_1X_1 + B_2X_2 = O$$

is satisfied on substitution from Equation 4.53.

$$(-B_1A_1^{-1}A_2 + B_2)X_2 = O \qquad (4.54)$$

According to the argument preceding Equation 4.45 the bracketed factor in Equation 4.54, is a null matrix, hence Equation 4.54 is satisfied whatever values were chosen for X_2.

The solution of the foregoing example was arbitrary to the extent that we were free to assign arbitrary values to the unknowns x_3 and x_4. In general it is possible to assign arbitrary values to as many unknowns as are in excess of the rank of the system. Thus if the rank of our example were $r = 3$, the non-singular submatrix A_1 would be of order 3×3, and we would be free to assign an arbitrary value to x_4 only.

It now remains to deal with the extreme case where the rank of the system equals the number of unknowns. As this is only possible in the square case, or in cases having an excess of equations over the number of unknowns, we can partition the system as follows:

$$A\,X = O$$

$$\begin{bmatrix} A_1 \\ B_1 \end{bmatrix} X = 0 \qquad (4.55)$$

Here A_1 has been arranged to be non-singular. From Equation 4.55 we can write

$$A_1X = O, \quad B_1X = O$$

and since A_1 is nonsingular both sides of the first equation can be premultiplied by A_1^{-1}.

$$A_1^{-1} A_1X = X = O \qquad (4.56)$$

From this it follows that *when the rank of the system equals the number of unknowns, the latter must necessarily be zero.* In the square case this result becomes the familiar statement that if the determinant of the system does not vanish, only the trivial, vanishing solution is possible.

Summary of solution

1. Find the rank r of the system. If r is less than the number of unknowns a non-trivial solution exists.
2. Pick out r equations with at least one minor of order $r \times r$ not zero. Denote the minor by A_1.
3. Solve for r unknowns X_1 according to Equation 4. 53.

$$X_1 = -A_1^{-1} A_2 X_2$$

where the excess unknowns X_2 are given arbitrary values.

4. Remaining equations are automatically satisfied.

4. 6 NON-HOMOGENEOUS EQUATIONS—THEIR CONSISTENCY AND SOLUTION

In Chapter 2 it was explained how to solve by matrix methods a system of non-homogeneous linear equations in the simplest case. This is the square case, in which the number of unknowns equals the number of equations, and the matrix of the system is nonsingular, so that its reciprocal can be formed.

We are now sufficiently equipped to tackle the general case, in which the number of unknowns does not equal the number of equations, and the matrix of the system is rectangular. We shall follow a similar line as we did with homogeneous equations; an example will be written out in full, and in parallel the general case will be explained. We take as an example 3 equations in 4 unknowns, and assume that the rank of the system is $r = 2$.

$$\begin{array}{l|l}
a_{11}x_1 + a_{12}x_2 & + a_{13}x_3 + a_{14}x_4 = h_1 \\
a_{21}x_1 + a_{22}x_2 & + a_{23}x_3 + a_{24}x_4 = h_2 \\
\hdashline
a_{31}x_1 + a_{32}x_2 & + a_{33}x_3 + a_{34}x_4 = h_3
\end{array}$$

$$AX = H \tag{4.57}$$

The above equations are arranged to have a non-vanishing minor of order $r = 2$ in the leading position. The general form of the equations is now partitioned according to the scheme

$$\begin{bmatrix} A_1 & A_2 \\ B_1 & B_2 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} \tag{4.58}$$

which corresponds to the dotted lines in Equations 4. 57. Applying a suitable elementary operation to the rows of A the submatrices B_1 and B_2 can be removed (see p. 132).

$$PAX = PH = G$$

$$\begin{bmatrix} A_1 & A_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} \tag{4.59}$$

On Equation 4. 57 this operation takes the form of adding suitable multiples of the first two equations to the third, with the result that the left hand side of the latter vanishes. Denoting the multipliers by c_1 and c_2 we obtain

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 &= h_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 &= h_2 \\
0 &= h_3 + c_1h_1 + c_2h_2
\end{aligned} \tag{4.60}$$

Using general symbols, Equations 4. 60 have the form

$$\begin{aligned}
A_1X_1 + A_2X_2 &= G_1 \\
O &= G_2
\end{aligned} \tag{4.61}$$

Before we attempt to solve the equations we must consider the question of their *consistency or compatibility*. Unless the column vector of constants on the right hand sides of Equations 4. 57 is such that the third of Equations 4. 60, or the second of Equations 4. 61 is automatically satisfied, the equations are inconsistent or incompatible.

A test for the consistency of a system of equations is found by comparing two matrices.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & h_1 \\ a_{21} & a_{22} & a_{23} & a_{24} & h_2 \\ 0 & 0 & 0 & 0 & (h_3 - c_1h_1 - c_2h_2) \end{bmatrix}$$

$$\begin{bmatrix} A_1 & A_2 \\ 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} A_1 & A_2 & G_1 \\ 0 & 0 & G_2 \end{bmatrix} \tag{4.62}$$

The right hand matrix is called the *augmented* matrix of the system. From (4. 62) we can see the condition of compatibility of a system of equations. Since for consistency the submatrix G_2 (which equals $h_3 + c_1h_1 + c_2h_2$) must vanish, it follows that the rank of the augmented matrix must be the same as the rank of the matrix of the equations. As the partially reduced matrices (4. 62) have been obtained from the original matrices of the system by elementary operations which do not alter rank, the foregoing condition can be expressed as follows

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}, \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & h_1 \\ a_{21} & a_{22} & a_{23} & a_{24} & h_2 \\ a_{31} & a_{32} & a_{33} & a_{34} & h_3 \end{bmatrix} \tag{4.63}$$

$$A, \quad [A\ H]$$

A system of equations is consistent if the rank of its matrix and the augmented matrix is the same

Assuming now that the given equations are compatible we can go on to solve them. Since in our example the third equation assumes the trivial form O = O we are left with only two to consider. The system is rewritten in the form

$$a_{11}x_1 + a_{12}x_2 = -a_{13}x_3 - a_{14}x_4 + h_1$$
$$a_{21}x_1 + a_{22}x_2 = -a_{23}x_3 - a_{24}x_4 + h_2$$
$$A_1X_1 = -A_2X_2 + G_1 \tag{4.64}$$

The unknowns x_3 and x_4 can now be given arbitrary values. This done, Equations 4.64 assume the form

$$a_{11}x_1 + a_{12}x_2 = f_1$$
$$a_{21}x_1 + a_{22}x_2 = f_2$$
$$A_1X_1 = F_1 \tag{4.65}$$

where A_1 is non-singular. The final solution is

$$X_1 = A_1^{-1}F_1 \tag{4.66}$$

by the basic method of Chapter 2.

The values of the unknowns thus obtained clearly satisfy the first r (= 2) equations of the system. That they also automatically satisfy the remaining m − r (= 1) equations can be verified in the same way as was done for homogeneous equations in the preceding section.

It is instructive to consider what happens when the values of x_3 and x_4 in Equations 4.64, are so chosen that the right hand side of those equations reduces to zero. We are then left with a set of 2 homogeneous equations in 2 unknowns and rank 2. By Equation 4.56 the only possible solution is $x_1 = x_2 = 0$. However, together with the previously chosen (finite) values of x_3 and x_4 we have, in fact, a non-trivial solution of the original system (Equation 4.57).

Summary of solution

1. Check for consistency by comparing the rank of the matrix of the system with the rank of the augmented matrix. The rank r must be the same for both matrices.
2. Pick out r equations with at least one minor order of r × r not zero. Denote it by A_1.

3. Solve for r unknowns X_1 according to Equation 4.66.

   $X_1 = A_1^{-1}F_1$

   where the excess unknowns X_2 are given arbitrary values.
4. Remaining equations are automatically satisfied.

4.7 LINEAR DEPENDENCE

Whenever we deal with a set of algebraic entities such as scalar numbers, functions, vectors, or matrices, we can attempt to form what is known as a relation of *linear dependence* between them. By this we mean an equation of the form

$$c_1f_1 + c_2f_2 + c_3f_3 + \cdots + c_nf_n = 0 \tag{4.67}$$

where the c-s are scalar constants and the f-s are, say, column vectors or some other algebraic entities under consideration.

If the Equation 4.67 is possible with at least one of the c-s *not zero*, the f-s are said to be *linearly dependent*. If all the c-s have to be put equal to zero to satisfy the above equation, the f-s are said to be *linearly independent*.

If the f-s happen to be linearly dependent, it is always possible to express one of them in terms of the others by an equation of the form

$$f_i = \sum_{j \neq i} k_j f_j \tag{4.68}$$

Written in this way the f_i is said to be a *linear combination* of the remaining f_j.

As an illustration of the foregoing concepts we write out in full the example of 3 column vectors, each containing 3 elements. A relation of linear dependence between them assumes the form

$$c_1f_1 + c_2f_2 + c_3f_3 = 0$$

$$c_1\begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} + c_2\begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} + c_3\begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix} = 0 \tag{4.69}$$

If the column vectors can be made to satisfy the above equation, with at least one of the constants c_j not equal to zero, they are said to be linearly dependent. The c_j are called constants of linear dependence.

Assuming that this is the case and that, say, c_2 does not vanish, the equation can be rewritten in the form

$$\begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} = -\frac{c_3}{c_2}\begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix} - \frac{c_1}{c_2}\begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} \tag{4.70}$$

which is analogous to Equation 4.68. Written in this way the column vector on the left is expressed as a linear combination of the remaining vectors.

The above example of linear dependence between column vectors can be put into matrix form, and when this is done we find that it takes us straight into the problem of solving homogeneous linear equations. Multiplying the vectors in Equation 4.69 by their associated constants and adding them matrixwise, we obtain the following matrix relation:

$$\begin{bmatrix} c_1a_{11} + c_2a_{12} + c_3a_{13} \\ c_1a_{21} + c_2a_{22} + c_3a_{23} \\ c_1a_{31} + c_2a_{32} + c_3a_{33} \end{bmatrix} = O$$

The reader will recognise that this is equivalent to the form

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = O$$

$$A\ C = O \tag{4.71}$$

which is that of a system of homogeneous linear equations, with the c_j in the role of unknowns.

Relations of linear dependence may exist between row vectors as well as column vectors. The possibility is shown by the following example of 3 row vectors, each having 3 elements.

$$l_1\begin{bmatrix} a_{11} & a_{12} & a_{13} \end{bmatrix} + l_2\begin{bmatrix} a_{21} & a_{22} & a_{23} \end{bmatrix} + l_3\begin{bmatrix} a_{31} & a_{32} & a_{33} \end{bmatrix} = O \tag{4.72}$$

The l-s are constants of linear dependence. If at least one of them does not vanish, the vectors are linearly dependent. Equation 4.72 can be put in matrix form as follows:

$$\begin{bmatrix} l_1 & l_2 & l_3 \end{bmatrix}\begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} = O$$

$$L\ A = O \tag{4.73}$$

Transposition of Equation 4.73 makes it clear that it is of the same basic form as Equation 4.71.

Equation 4.71 gives the relation of linear dependence in matrix form in the special case where the number of vectors is the same as the number of elements in each vector. In this case the vectors form a square matrix when assembled together. Relations of linear dependence can be formed in more general cases, in which the number of vectors does not equal the number of elements in each vector. The matrices obtained in such cases are rectangular, as in the following example of 4 vectors, each having 3 elements.

$$c_1\begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \end{bmatrix} + c_2\begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} + c_3\begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix} + c_4\begin{bmatrix} a_{14} \\ a_{24} \\ a_{34} \end{bmatrix} = O \tag{4.74}$$

Matrix form of Equation 4.74:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = O$$

$$A\ C = O \tag{4.75}$$

With the help of our knowledge of homogeneous equations we are now in a position to state the general condition under which a set of vectors is linearly dependent. Since the constants of linear dependence can be considered to be the unknowns in a set of equations, they can be given non-zero values if the rank of the system is less than the number of constants. *In other words, the matrix made up of the given vectors must be of lower rank than the number of vectors*. This statement applies to both columns and rows.

Expressed in terms of minors the above condition means that all minors of order equal to, or greater than the number of linearly independent vectors must vanish. In the square case, linear dependence of both rows and columns implies the vanishing of the determinant of the associated matrix.

In the above example, Equation 4.75, the linear dependence of the 4 columns is assured, since the rank of the matrix can be at most $r = 3$. Hence Equation 4.74 can be written down always, with at least one of the c-s not equal to zero.

The 3 rows of the matrix on the left hand side of Equation 4.75 will be linearly dependent only if the rank is $r \leqslant 2$. If $r = 3$ the rows will be linearly independent.

In the square case, exemplified by Equations 4.69 and 4.71, the condition of linear dependence for both rows and columns is that the rank of the square

matrix must be less than its order. In other words the determinant of the matrix must vanish. The associated system of homogeneous equations then has a nontrivial solution, and a non-zero set of constants of linear dependence can be found. It should be noted that the condition of vanishing determinant is implied by Equations 4.69 and 4.70 and Rule 5 of determinants. The relation of linear dependence represented by Equation 4.70 means that when suitable multiples of the first and third columns are subtracted from the second column, the latter will vanish.

It follows from the foregoing discussion that the rank of a matrix and the linear dependence of its rows and columns are largely interchangeable concepts. When we say that a matrix is of *rank* r, we imply that (r + 1) *rows or columns* of this matrix will be found to be *linearly dependent*, or that we can find r *linearly independent rows or columns* in the matrix. Conversely, a matrix which contains r linearly independent rows or columns is of rank r.

The terminology of linear dependence is frequently applied to systems of equations. When it is stated that a given set of homogeneous equations is linearly dependent, it is meant that the rows of its matrix are linearly dependent. As this implies that the rank of the system is less than the number of equations, the excess of equations over rank can be got rid of with the help of an elementary operation as in Equations 4.50 or 4.51. If, moreover, the number of unknowns is greater than the rank of the system, we are assured of a nontrivial solution by the method of Section 4.5. When a set of non-homogeneous equations is said to be linearly dependent it is meant that both sides of the equations are linearly dependent. This implies that the rows of the augmented matrix are linearly dependent. It should be borne in mind, however, that this does not necessarily mean that the equations are consistent and soluble.

The concepts of linear dependence of vectors and rank of matrices are very important and frequently used. The reader is advised to refer back to this section, and the section dealing with rank, until he has fully grasped the meaning and implications of these terms.

4.8 THE EIGENVALUE PROBLEM ASSOCIATED WITH A SQUARE MATRIX

When the equivalence of matrices was explained in an earlier section it was found that a given matrix could be reduced to diagonal form in a variety of ways, each resulting in different values for the diagonal elements. The differences between the individual reductions resulted from our freedom to carry out the requisite elementary operations on rows and columns in different sequences. In this section we propose to introduce an altogether unique method of diagonalising a square matrix, which always yields the same diagonal form. Since the subject comes up in many important applications it deserves close study.

To approach the problem we go back to linear transformations as introduced

in the opening section of Chapter 2. There we found that a linear transformation of the form

$$X^{(1)} = AX \tag{4.76}$$

could be visualised as a transformation of points, or point transformation, in a fixed system of coordinates. Here we note that a point transformation may be conceived as a transformation of vectors, which radiate from the origin and have components equal to the coordinates of points. Thus the transformation (Equation 4.76) may be represented in 2 dimensions by Fig. 4.1.

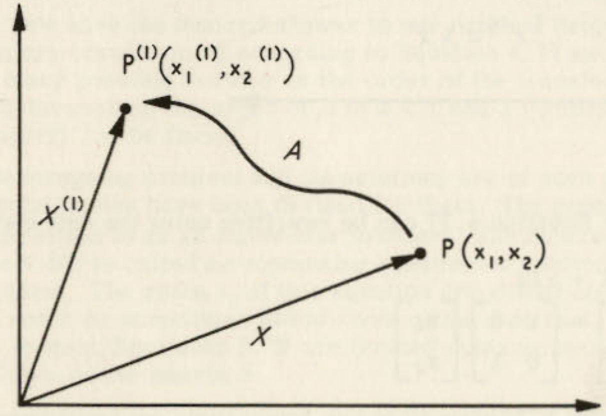


Fig. 4.1

We now ask the question whether it is possible to find a vector X, which will be transformed by a given square matrix A into one of the form $X^{(1)} = \lambda X$. It should be noted that the new vector has the same direction as the original one, but differs from it in length (see Fig. 4.2). We shall find that the answer to our question is in the affirmative, that usually a number of such vectors exists, and when trying to find them we shall also discover a method of diagonalising the square matrix A.

We shall solve the problem in general terms but in parallel we shall also write out the special case of order 2 × 2 as an illustration.

The question is formulated by the equation

$$AX = \lambda X$$

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{4.77}$$

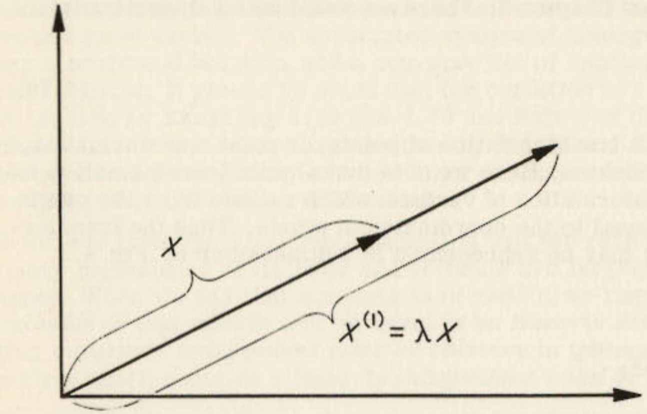


Fig. 4.2

The right hand side of Equation 4.77 can be rewritten using the unit matrix.

$$AX = \lambda IX$$

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{4.78}$$

Transferring the right hand side to the left we obtain a system of homogeneous equations with a square matrix.

$$(A - \lambda I)X = O$$

$$\begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = O \tag{4.79}$$

By Section 4.5, Equation 4.79 can be solved if its rank is less than its order $r < n$ or, in other words, if the determinant of the system vanishes.

$$|A - \lambda I| = 0$$

$$\begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = O \tag{4.80}$$

The above condition constitutes an algebraic equation of degree n (2 in our example). In general we expect n solutions, or n values of the multiplier λ, each of which can be substituted into the system of Equations 4.79. Each

value of λ thus found ensures the vanishing of the determinant of the Equations 4.79 and hence permits their solution.

Denoting the roots of the determinantal Equation 4.80 by λ_i we substitute them in turn into the Equations 4.79 and find the corresponding vectors of solutions X_i by the methods of Section 4.5.

$$(A - \lambda_i I) X_i = 0, \qquad AX_i = \lambda_i X_i$$

$$\begin{bmatrix} a_{11} - \lambda_i & a_{12} \\ a_{21} & a_{22} - \lambda_i \end{bmatrix} \begin{bmatrix} x_{1i} \\ x_{2i} \end{bmatrix} = O = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_{1i} \\ x_{2i} \end{bmatrix} = \begin{bmatrix} \lambda_i x_{1i} \\ \lambda_i x_{2i} \end{bmatrix} \tag{4.81}$$

We now have the desired answer to our original problem Equation 4.77. Vectors transforming according to Equation 4.77 exist. There are in general as many possible vectors as the order of the transforming matrix A. Thus in 2 dimensions the order of A is 2×2, and 2 vectors having the required property can be found.

The foregoing problem and its solutions are of such great importance that special names have been devised for them. The problem (Equation 4.77) is referred to as an *eigenvalue problem*. The determinantal equation (Equation 4.80) is called an *eigenvalue equation* or *characteristic equation* of the problem. The roots λ_i of this equation are called *eigenvalues*, or *characteristic roots*, or sometimes *latent roots of the matrix A*. The solutions X_i of the system, Equations 4.79 are termed *eigenvectors*, or *characteristic vectors*, of the matrix A.

Having found the eigenvalues and eigenvectors of the square matrix A, we can now use them to diagonalise A, as was suggested in the opening paragraphs of this section. As a first step we collect all the n eigenvectors X_i, multiply them by the corresponding eigenvalues λ_i, and assemble them into a square matrix of order $n \times n$.

$$\left[\lambda_1 X_1 \lambda_2 X_2 \ldots \lambda_i X_i \ldots \lambda_n X_n\right]$$

$$\begin{bmatrix} \lambda_1 x_{11} & \lambda_2 x_{12} \\ \lambda_1 x_{21} & \lambda_2 x_{22} \end{bmatrix} \tag{4.82}$$

Next we observe that (4.82) can be rewritten as a product of two matrices as follows:

$$\begin{bmatrix} \lambda_1 x_{11} & \lambda_2 x_{12} \\ \lambda_1 x_{21} & \lambda_2 x_{22} \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = Q\Lambda$$

$$\left[\lambda_1 X_1 \ldots \lambda_n X_n\right] = \begin{bmatrix} x_{11}\, x_{12} \ldots x_{1n} \\ x_{21}\, x_{22} \ldots x_{2n} \\ \ldots\ldots\ldots\ldots \\ x_{n1}\, x_{n2} \ldots x_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1\, 0 \ldots 0 \\ 0\, \lambda_2 \ldots 0 \\ \ldots\ldots\ldots \\ 0 \ldots\ldots \lambda_n \end{bmatrix} = Q\Lambda \tag{4.83}$$

Q is the square matrix made up of the eigenvectors of A placed side by side. Λ is the diagonal matrix having the eigenvalues of A for its diagonal elements. The product $Q\Lambda$ thus contains all possible values of the *right hand side of Equation 4.81*, after the eigenvalue problem has been solved.

Likewise we can collect all possible values of the *left hand side of Equation 4.81* into a single square matrix of order $n \times n$, and then express it as a product of two square matrices.

$$[AX_1\ AX_2\ \ldots.\ AX_i\ \ldots.\ AX_n] = AQ$$

$$\left[\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix} \quad \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_{12} \\ x_{22} \end{bmatrix}\right] = \tag{4.84}$$

$$\begin{bmatrix} a_{11}x_{11} + a_{12}x_{21} & a_{11}x_{12} + a_{12}x_{22} \\ a_{21}x_{11} + a_{22}x_{21} & a_{21}x_{12} + a_{22}x_{22} \end{bmatrix} = AQ$$

As a final step in the diagonalisation procedure we equate the Expressions 4.83 and 4.84, as we can do, by virtue of Equation 4.81.

$$AQ = Q\Lambda$$

$$[AX_1\ AX_2\ \ldots\ AX_i\ \ldots\ AX_n] = [\lambda_1 X_1\quad \lambda_2 X_2\ \ldots\ \lambda_i X_i\ \ldots\ \lambda_n X_n] \tag{4.85}$$

Premultiplication of both sides of Equation 4.85 by Q^{-1} yields the result

$$Q^{-1}AQ = \Lambda$$

$$\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}^{-1}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \tag{4.86}$$

According to Equation 4.86 a square matrix is reduced to a diagonal form having its own eigenvalues as diagonal elements. The result is obtained through the use of one matrix Q, made up of the eigenvectors of A. The diagonal form Equation 4.86 is sometimes called the *canonical* form of the square matrix A. The reduction $Q^{-1}AQ$ is referred to as a *similarity transformation*.

The diagonalisation procedure outlined above is subject to certain limitations. In the first place the eigenvalue equation, Equation 4.80, may not have the full complement of n distinct roots. In such cases difficulties may be encountered in constructing a nonsingular transformation matrix Q from the eigenvectors of A, and hence in working out the diagonal form Λ. It is beyond the scope of the present treatment to go into these difficulties, but one general rule will be stated without proof. If the n eigenvalues are distinct, it can be shown that the corresponding eigenvectors are linearly independent. Hence the reducing matrix Q is of rank n and therefore nonsingular. The reduction is then always possible.

Finally it should be noted that the reducing matrix Q can be built up of row eigenvectors of A instead of columns. In such cases the eigenvalue problem (Equation 4.77) is formulated as follows:

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \lambda\begin{bmatrix} x_1 & x_2 \end{bmatrix}$$

As there is no basic difference between this set of equations and Equation 4.77 the procedure follows the same lines as above.

Equation 4.86 can be given an alternative interpretation to the one explained above. Rewriting it in the form

$$A = Q\Lambda Q^{-1} \tag{4.87}$$

we see that the *matrix A has been decomposed* into a product of three factors. The factors are of a special form, as they include the diagonal form of A and the matrix Q, which consists of eigenvectors of A. When discussing applications in the chapters to follow, we shall find the point of view embodied by Equation 4.87 very useful. It will help to resolve several difficult problems.

The geometrical meaning of the eigenvalue problem mentioned at the beginning of this section lends itself to an instructive extension. A transformation of points, or radius vectors, represented by the matrix A in one system of co-ordinates will no longer be represented by this matrix if the co-ordinates are changed. To find the form of the transformation matrix in the new co-ordinates let us write

$$Y = RX \tag{4.88}$$

where Y represents the new co-ordinates of the original point P. Similarly the co-ordinates of the transformed point $P^{(1)}$ in the new frame of reference are

$$Y^{(1)} = RX^{(1)} \tag{4.89}$$

The points are best thought of as fixed to the sheet of paper in Fig. 4.3, while the co-ordinates OX_1X_2 are transformed into OY_1Y_2. Let us now substitute the above expressions into the transformation

$$X^{(1)} = AX$$

We find

$$R^{-1}Y^{(1)} = AR^{-1}Y$$

Premultiplication of both sides by R yields

$$Y^{(1)} = RAR^{-1}Y = BY \tag{4.90}$$

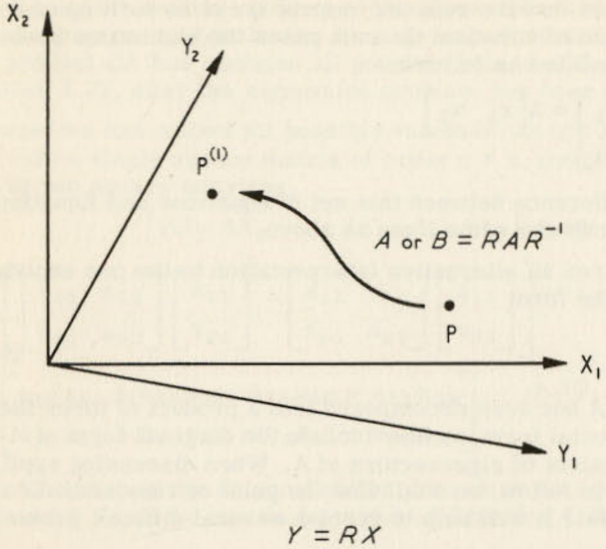


Fig. 4.3

Hence the matrix B, which transforms the point P into $P^{(1)}$ in the new system of co-ordinates, is obtained from the old matrix A through a similarity transformation. The similarity transformation involves the matrix R, which changes the original co-ordinates into the new ones.

The linear transformation A assumes a particularly simple form in a frame of reference obtained through the use of the matrix Q Letting

$$R = Q^{-1}, \quad Y = Q^{-1}X \tag{4.91}$$

we obtain

$$Y^{(1)} = Q^{-1}AQY = \Lambda Y$$

or

$$y_i^{(1)} = \lambda_i y_i$$

Hence the problem of transforming the point P into the point $P^{(1)}$ in this particular frame of reference reduces to scalar multiplication of each co-ordinate by the corresponding eigenvalue of the matrix A.

The geometrical relation between the two co-ordinate systems used above is such that the eigenvectors of A in the X-system correspond to unit or base

vectors in the Y-system. To see this we take the reciprocal of Equations 4.91 and write it out in full for the unit vector along the y_1-axis.

$$X = QY$$

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_{11}\,x_{12} \ldots x_{1n} \\ x_{21}\,x_{22} \ldots x_{2n} \\ \ldots\ldots\ldots\ldots \\ x_{n1}\,x_{n2} \cdots x_{nn} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix} = X_1$$

The foregoing geometrical example affords a good illustration of the advantages to be derived from choosing a convenient co-ordinate system in problems of applied mathematics. When discussing electrical applications we shall have ample opportunity to use the methods outlined in this section.

4.9 UNITARY MATRICES AND ORTHOGONALITY

The concepts and methods introduced in the preceding sections will now be applied to some special types of matrices which are of importance in applications.

We recall that linear transformations were introduced in Chapter 2 on the example of rotations of co-ordinates. In two dimensions a rotation is represented by the equation

$$\begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{4.92}$$

The matrix of this transformation provides an example of an important class of matrices called orthogonal. The outstanding property of an *orthogonal matrix, always real*, is the fact that its transpose equals its reciprocal. In symbols this is stated as follows:

$$A' = A^{-1}$$

$$AA' = A'A = I \tag{4.93}$$

The reader will check that the matrix (right hand side of Equation 4.92) satisfies this relation, having first observed that its determinant equals unity.

The reason for the name orthogonal matrix derives partly from the fact that when used as a transformation matrix it transforms a rectangular co-ordinate system into another rectangular system, as exemplified by rotations about the origin. Another reason for the name is to be found in the properties of the row and column vectors that make up an orthogonal matrix.

Taking, say, two column vectors $[a_{is}]$ and $[a_{it}]$, where s and t are fixed subscripts, from an orthogonal matrix A, and forming the product

$$\left[a_{is}\right]' \left[a_{it}\right] \tag{4.94}$$

we find by Equation 4.93 that it equals either unity or zero. This is so because the product is an element of the matrix product $A'A$. It vanishes when it is an off diagonal element, in which case it is composed of 2 different vectors, and it equals unity when it is a diagonal element, in which case it is composed of one and the same vector. Using the Kronecker δ-symbol these statements can be summarised by the equation

$$\left[a_{is}\right]' \left[a_{it}\right] = \delta_{st} \tag{4.95}$$

Expressed in terms of summations, Equation 4.95 assumes the form

$$\sum a_{is}\, a_{it} = \delta_{st} \tag{4.96}$$

Vectors having the above property are said to form *orthogonal sets of vectors*. We thus see that the rows or columns of orthogonal matrices form orthogonal sets of vectors. Orthogonality of real vectors has a simple geometrical or physical meaning. Vectors representing displacements, velocities, or forces in ordinary three or two dimensional space are orthogonal, when their directions are perpendicular to each other. The rows of the matrix of Equation 4.92 can be taken to represent a pair of vectors in the plane OXY, and a moments reflection will satisfy the reader that they are at right angles to each other.

Products of vectors of the form discussed above are called *scalar products*, because they yield a scalar number.

It is an instructive example in matrix methods to show that a symmetric matrix is reduced to canonical diagonal form by an orthogonal matrix Q. Thus, Equation 4.86 assumes the following form when A is symmetric

$$Q^{-1}AQ = Q'AQ \tag{4.97}$$

To demonstrate this, all we have to do is to show that the eigenvectors X_i, which make up the matrix Q in this case, satisfy the relation

$$X_s'X_t = \delta_{st}$$

$$Q'Q = I \tag{4.98}$$

To this end we rewrite Equation 4.81 for the eigenvector X_s in transposed form

$$\lambda_s X_s' = X_s'A' = X_s'A \tag{4.99}$$

where the right hand side follows by virtue of the symmetry of A.
Further we write down Equation 4.81 for the eigenvector X_t and premultiply both sides by A^{-1}.

$$\lambda_t X_t = AX_t$$

$$\lambda_t A^{-1}X_t = X_t \tag{4.100}$$

Finally we premultiply Equation 4.100 by $\lambda_s X_s'$ and find with the help of Equation 4.99

$$\lambda_t X_s'AA^{-1}X_t = \lambda_s X_s'X_t$$

$$\lambda_t X_s'X_t = \lambda_s X_s'X_t$$

$$(\lambda_s - \lambda_t)\, X_s'X_t = 0 \tag{4.101}$$

Since, in general, we assume that the eigenvalues of A are distinct, the above equation can be satisfied only if

$$X_s'X_t = 0 \tag{4.102}$$

This result shows that the off diagonal elements of the product $Q'Q$ vanish. It remains to show that the diagonal elements do not vanish and can be made equal to unity.

To prove this we recall that the eigenvectors X_i are non-trivial solutions of the system of homogeneous equations (Equations 4.79), and therefore each has at least one non-zero element. Hence the product $X_s'X_s$ does not vanish.

$$X_s'X_s \neq 0 \tag{4.103}$$

Moreover, each eigenvector is arbitrary at least to the extent of a multiplicative factor, by Equation 4.53. Choosing this factor suitably makes it possible to give the scalar product (4.103) the value unity.

$$X_s'X_s = 1 \tag{4.104}$$

This establishes the orthogonality of the matrix Q, and proves Equation 4.97 for a symmetric matrix A.

Orthogonal matrices are a special case of a more general class of *complex matrices* called unitary. The defining property of a *unitary matrix* is

$$\bar{A}' = A^{-1} \text{ or } \bar{A}'A = I \tag{4.105}$$

In words, the Hermitian adjoint of a unitary matrix equals its reciprocal.

Example:

$$A = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 & 1+j \\ 1-j & -1 \end{bmatrix} \tag{4.106}$$

The reader is invited to write down the Hermitian adjoint and the reciprocal of this matrix and satisfy himself that it is unitary.

The rows and columns of a unitary matrix can be said to form orthogonal sets of vectors, if a generalised definition of the scalar product is adopted. Instead of Equation 4.94 the scalar product of complex vectors is defined to be

$$\overline{\begin{bmatrix} a_{is} \end{bmatrix}}' \begin{bmatrix} a_{it} \end{bmatrix} = \sum_i \overline{a_{is}} a_{it} \tag{4.107}$$

Following this definition it can be seen from Equations 4.105 that the rows and columns of unitary matrices are orthogonal because they satisfy the relation

$$\overline{\begin{bmatrix} a_{is} \end{bmatrix}}' \begin{bmatrix} a_{it} \end{bmatrix} = \sum_i \overline{a_{is}} a_{it} = \delta_{st} \tag{4.108}$$

Unitary matrices are a generalised form of orthogonal matrices in a similar way as Hermitian matrices are a more general form of symmetric matrices. Hence we may expect that a Hermitian matrix is diagonalised by a unitary matrix according to a suitable modification of Equation 4.97.

$$Q^{-1} HQ = \overline{Q}' HQ \tag{4.109}$$

Equation 4.109 is proved in the same way as Equation 4.97, except that the generalised definition of the scalar product, Equation 4.107, must be used.

4.10 DIFFERENTIATION AND INTEGRATION OF MATRICES

Matrices whose elements are functions of one or more independent variables are frequently encountered in applications. In the work to follow we shall have occasion to deal with vectors whose elements are functions of time, and we shall find it necessary to differentiate and integrate them. In the present section we introduce the notation and definitions customarily used in this connection.

The functional dependence of a matrix is denoted in the same way as that of an ordinary function. Thus, if all elements of, say, a column vector are functions of time we write

$$X = X(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} \tag{4.110}$$

The various circuit matrices discussed in Chapter 3 are in general functions of frequency. To emphasise its dependence on frequency the admittance matrix of a two-port network may be written as follows:

$$Y(\omega) = \begin{bmatrix} y_{11}(\omega) & y_{12}(\omega) \\ y_{21}(\omega) & y_{22}(\omega) \end{bmatrix} \tag{4.111}$$

Differentiation of matrices of the type of Equations 4.110 or 4.111 means the differentiation of each element with respect to the independent variable. Thus, the first time derivative of the vector on the right hand side of Equation 4.110, is

$$\frac{d}{dt}X(t) = \begin{bmatrix} \frac{d}{dt}x_1(t) \\ \frac{d}{dt}x_2(t) \\ \frac{d}{dt}x_3(t) \end{bmatrix} = \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} = \dot{X}(t) \tag{4.112}$$

As in the case of ordinary functions, the functional notation is dropped whenever it is clearly implied by the context. Equation 4.112 then simplifies to the form

$$\frac{dX}{dt} = \dot{X} \tag{4.113}$$

From the foregoing definition it follows that the differential operator can be treated as a scalar factor premultiplying the matrix, element by element, as laid down in Chapter 2, and in full analogy to differential operators applied to ordinary functions.

The differentiation rule, Equation 4.112, can be given an obvious kinematic interpretation. If the column $X(t)$ represents the position vector of a point as a function of time, its first derivative $\dot{X}(t)$ is the velocity vector of the point. The second derivative $\ddot{X}(t)$ is then the acceleration of the point.

Integration of matrices follows the same pattern as differentiation, in that the

integration operator can be treated as a scalar factor multiplying all matrix elements. Thus we write by analogy with Equations 4.112

$$\int X(t)dt = \begin{bmatrix} \int x_1(t)dt \\ \int x_2(t)dt \\ \int x_3(t)dt \end{bmatrix} = \int X dt \tag{4.114}$$

A more complicated example is

$$\int \begin{bmatrix} e^{-pt} & 0 \\ 0 & e^{-pt} \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} dt = \begin{bmatrix} \int e^{-pt} u_1(t)dt \\ \int e^{-pt} u_2(t)dt \end{bmatrix} \tag{4.115}$$

Integrals of the form of Equation 4.115, will be encountered in connection with the Laplace transformation.

4.11 FUNCTION OF A MATRIX

The multiplication rule of matrix algebra allows us to multiply a square matrix by itself repeatedly. In this way any integral power of a square matrix may be formed. For example, given the matrix A, the expression A^n can always be worked out.

By analogy with the scalar power x^n we can regard A^n as a simple function of the matrix A, which will assume a variety of 'values', according to what numerical elements are inserted into A.

Drawing on the operations of multiplication by a scalar and addition of matrices, we can go on to form more complicated functions of matrices, analogous to scalar polynomials. Thus the expression

$$f(A) = aA^3 + bA^2 + cA \tag{4.116}$$

describes a well defined square matrix $f(A)$, obtained from A according to the prescription on the right hand side of Equation 4.116. Substitution of a variety of elements into A will yield a variety of matrices $f(A)$, which can be called the values of the function.

A natural extension of a polynomial function is an infinite power series. Since the formation of a power series requires the operations of multiplication and addition only, we can form power series of matrices. Thus, by analogy with the scalar exponential function

$$e^x = 1 + \frac{1}{1!}x + \frac{1}{2!}x^2 + \dots + \frac{1}{n!}x^n + \dots$$

we can form a function of the matrix A.

$$f(A) = e^A = I + \frac{1}{1!}A + \frac{1}{2!}A^2 + \dots + \frac{1}{n!}A^n + \dots \tag{4.117}$$

The above example suggests that we take as the definition of a *function of a matrix*, $f(A)$, the Taylor series expansion of f. We recall that the Taylor series of a scalar function $f(x)$ is

$$f(x) = \sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{d^n f}{dx^n}\right)_{x=a} (x-a)^n \tag{4.118}$$

and we define $f(A)$ to be the matrix

$$f(A) = \sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{d^n f}{dx^n}\right)_{x=a} (A-aI)^n \tag{4.119}$$

The power series definition of a function of a matrix assumes a relatively simple form when the matrix is diagonal. As an example let us write down the exponential function of the matrix

$$A = D = \begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix}$$

$$f(D) = e^D =$$

$$\begin{bmatrix} 1 + d_1 + \frac{1}{2!}d_1^2 + \dots & 0 & 0 \\ 0 & 1 + d_2 + \frac{1}{2!}d_2^2 + \dots & 0 \\ 0 & 0 & 1 + d_3 + \frac{1}{2!}d_3^2 + \dots \end{bmatrix}$$

$$e^D = \begin{bmatrix} e^{d_1} & 0 & 0 \\ 0 & e^{d_2} & 0 \\ 0 & 0 & e^{d_3} \end{bmatrix} \tag{4.120}$$

The function e^D consists in effect of three distinct scalar exponential functions.

The above example is typical of all functions of diagonal matrices. They consist of as many scalar functions of the same form as there are diagonal elements.

We are now in a position to introduce an important generalisation of the eigenvalue relation, Equation 4.77. We shall show that if

$$AX = \lambda X$$

then

$$f(A)X = f(\lambda)X \tag{4.121}$$

where $f(A)$ is any function of the matrix A defined by Equation 4.119 above. In the first place we observe that

$$A^2X = AAX = A\lambda X = \lambda AX = \lambda^2 X \tag{4.122}$$

(see p. 34 Chapter 2). Equation 4.122 can be extended step by step to any power A^n. Next we write

$$(A^r + A^s)X = A^rX + A^sX = (\lambda^r + \lambda^s)X \tag{4.123}$$

Again this can be seen to apply to any number of terms. Finally we form the relation

$$f(A)X = \sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{d^n f}{dx^n}\right)_{x=a} (A - aI)^n X \tag{4.124}$$

Expanding the bracketed expression $(A - aI)^n$ in powers of A, applying Equations 4.122 and 4.123 and collecting the terms again we find

$$f(A)X = \sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{d^n f}{dx^n}\right)_{x=a} (\lambda - a)^n X$$

$$f(A)X = f(\lambda)X \tag{4.125}$$

Equation 4.125 shows that the operation of $f(A)$ *on an eigenvector of* A *is equivalent to multiplication by the scalar* $f(\lambda)$

The function $f(A)$ can be diagonalised by the same matrix that diagonalises A

Thus we can prove that

$$Q^{-1}f(A)Q = f(\Lambda)$$

$$f(A) = Qf(\Lambda)Q^{-1} \tag{4.126}$$

provided that

$$Q^{-1}AQ = \Lambda \quad \text{or} \quad AQ = Q\Lambda \tag{4.127}$$

Equation 4.126 is established by showing first that it holds for a power of A and then extending the result to the power series form of $f(A)$. We write

$$Q^{-1}A^nQ = Q^{-1}A^{(n-1)}AQ$$

$$= Q^{-1}A^{(n-1)}Q\Lambda$$

by Equation 4.127. Continuing the process step by step we arrive at the result

$$Q^{-1}A^nQ = Q^{-1}Q\Lambda^n = \Lambda^n \tag{4.128}$$

Expressions of the form $(A - aI)^n$ present no difficulty once they are expanded in powers of A. It is then found that

$$Q^{-1}A^r a^{(n-r)} I Q = \Lambda^r a^{(n-r)} I$$

by virtue of the fact that the scalar matrix $a^{(n-r)}I$ commutes with any matrix. Hence we find

$$Q^{-1}f(A)Q = Q^{-1}\left\{\sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{d^n f}{dx^n}\right)_{x=a} (A - aI)^n\right\}Q$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{d^n f}{dx^n}\right)_{x=a} (\Lambda - aI)^n$$

$$= f(\Lambda)$$

We shall see in subsequent work that Equations 4.125 and 4.126 can be used to simplify many complicated problems. As an immediate example let us see how they can be applied to the exponential function. By Equation 4.125, the operation of e^A on an eigenvector of A takes on the form

$$e^A X = e^{\lambda} X$$

The diagonal form of e^A is

$$Q^{-1}e^A Q = e^{\Lambda}$$

or the decomposition of e^A is

$$e^A = Qe^{\Lambda}Q^{-1}$$

The Taylor series definition of a function of a *square* matrix should not be confused with the expansion of elements of *any* matrix in their Taylor series. Given any matrix, say a column vector X, whose elements are functions of an independent variable, the following expansion can be readily written down.

$$X(t) = \sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{d^n}{dt^n}X(t)\right)_{t=a} (t - a)^n \tag{4.129}$$

The above expression should be considered to be no more than a condensed way of writing the Taylor series for all the individual elements of the vector $X(t)$. The powers $(t - a)^n$ are merely scalar factors multiplying the elements of the matrices $\left(\frac{d^n}{dt^n}X(t)\right)_{t=0}$ As such they should be contrasted with the factors $(A - aI)^n$ in Equation 4.119, which are square matrices.

Chapter 5

Differential Equations of Passive Linear Circuits

The importance of the algebraic concepts and methods introduced in the preceding chapter is far reaching, particularly in problems of applied mathematics. In this chapter we shall show how to use these methods when dealing with differential equations of passive linear circuits.

As in Chapter 4 no attempt is made to approach mathematical rigour. Instead, every effort is made to explain algebraic procedures clearly and concisely, and to relate them to physical results.

5.1 DIFFERENTIAL EQUATIONS OF THE FORCE FREE NETWORK

In the present section we shall set up the differential equations for the currents in a lumped linear circuit without any applied voltage sources. We shall start with the example of a one-loop network consisting of a capacitance, inductance and resistance connected in series, and then extend our discussion to a general n-loop network.

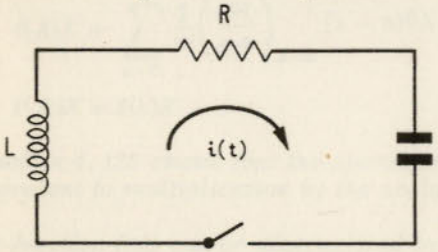


Fig. 5.1

Fig. 5.1 shows the basic one-loop circuit, which is the fundamental building block of any lumped network. We assume that the capacitor C has some charge on its plates, so that a fixed voltage exists across it initially. The problem that we set out to solve is to find the loop current i(t) as a function of time, from the moment the circuit is closed by depressing the key.

The instantaneous voltages across the three circuit elements are:

across L	$L\frac{di}{dt}$	
across R	Ri	
across C	$\frac{1}{C}\int i\,dt$	(5.1)

Applying Kirchoff's Voltage Law to the circuit after the key is depressed, we obtain the equation

$$L\frac{di}{dt} + Ri + \frac{1}{C}\int i\,dt = 0 \qquad (5.2)$$

Differentiating once with respect to time, a differential equation of the second order results.

$$L\,\frac{d^2i}{dt^2} + R\frac{di}{dt} + \frac{1}{C}i = 0 \qquad (5.3)$$

The right hand side of Equation 5.3 is zero because no external voltage source, or force, is applied to the circuit. For this reason Equation 5.3 is said to be homogeneous.

In this chapter we will not follow the standard method of solution of Equation 5.3, but will apply matrix methods. As a first step we introduce new symbols, which will prove better adapted to the application of matrix algebra. We let

$$i(t) = u_1(t) = u_1$$

$$L\dot{u}_1 = u_2 \qquad (5.4)$$

Equations 5.3 and 5.4 can now be rewritten in the form

$$\dot{u}_1 = 0.\,u_1 + \frac{1}{L}u_2$$

$$\dot{u}_2 = -\frac{1}{C}u_1 - \frac{R}{L}u_2$$

The single differential equation of the second order (Equation 5.3) has thus been replaced by two separate equations of the first order, at the cost of introducing a second dependent variable $u_2(t)$. Using matrix symbols, we have

$$\begin{bmatrix} \dot{u}_1 \\ \dot{u}_2 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{L} \\ -\frac{1}{C} & -\frac{R}{L} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \qquad (5.5)$$

$$\dot{U} = AU \quad \text{or} \quad \frac{d}{dt}U(t) = AU(t)$$

The differentiation of the column vector U with respect to time is carried out according to the rules explained in Section 4.10.

Equation 5.5 has been obtained on the assumption that L does not vanish. If it does, Equation 5.3 reduces to the simple form

$$\frac{di}{dt} = -\frac{1}{RC}\, i \tag{5.6}$$

which can be solved by direct integration. Written down in general symbols, Equation 5.6 can be considered a special case of Equation 5.5

$$\dot{u} = au \tag{5.7}$$

Here u(t) and a are scalars, or matrices of order 1×1.

In Equation 5.5 we have column vectors of 2 elements and a square matrix of order 2×2. It is customary to classify problems of the above type by saying that they have as *many degrees of freedom, or dimensions, or independent coordinates* as the order of their matrix. The column vector U can be used to define a space of two dimensions, in which the elements u_1 and u_2 would appear as units, or *base vectors*, along the co-ordinate axes. Adopting this point of view, the one loop network of Fig. 5.1 constitutes a dynamical problem with two independent coordinates, or degrees of freedom. The electrical interpretation of this fact is not difficult to see. To specify the electrical state of the circuit at any given instant it is necessary to give two items of information, e.g. the instantaneous current and the voltage across the capacitor. Alternatively it would be enough to state the current and its rate of change, which amounts to fixing the elements of the vector U. The two dimensional character of our problem also appears when the constants of integration are to be determined. The differential equation of the second order, Equation 5.3, or its equivalent equations of the first order, Equation 5.5, can only be fully solved after two constants have been stated.

Having set up the differential equations of the basic circuit in the matrix form of Equation 5.5, we postpone their solution until the next section, and in the meantime go on to formulate the equations of the general n-mesh network.

As a first step we write down the equations of the two loop network of Fig. 5.2. We omit the possibility of mutual inductance between the loops, as it complicates the detailed form of the problem without introducing any new principles. Application of Kirchoff's Law to the circuit, immediately after the key is depressed, yields the equations

$$\left[\left(L_1 + L_2\right)\frac{d}{dt} + R_1 + \left(\frac{1}{C_1} + \frac{1}{C_2}\right)\int dt\right] i_1 - \left[L_2\frac{d}{dt} + R_1 + \frac{1}{C_2}\int dt\right] i_2 = 0$$

$$-\left[L_2\frac{d}{dt} + R_1 + \frac{1}{C_2}\int dt\right] i_2 + \left[\left(L_2 + L_3\right)\frac{d}{dt} + \left(R_1 + R_2\right) + \left(\frac{1}{C_2} + \frac{1}{C_3}\right)\int dt\right] i_2 = 0 \tag{5.8}$$

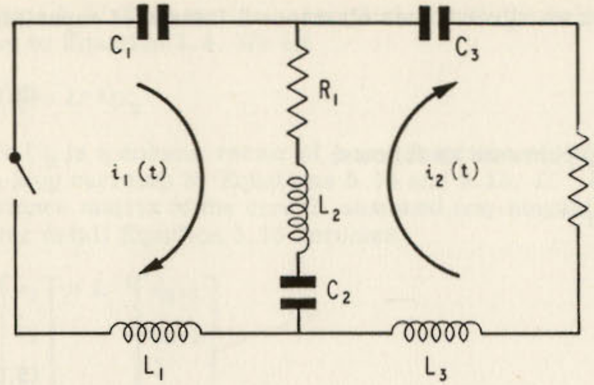


Fig. 5.2

Equations 5.8 are simplified by the use of double subscript symbols and by differentiation with respect to time.

$$\begin{aligned}
&\left(L_{11}\frac{d^2}{dt^2} + R_{11}\frac{d}{dt} + \frac{1}{C_{11}}\right)i_1 + \left(L_{12}\frac{d^2}{dt^2} + R_{12}\frac{d}{dt} + \frac{1}{C_{12}}\right)i_2 = 0\\
&\left(L_{21}\frac{d^2}{dt^2} + R_{21}\frac{d}{dt} + \frac{1}{C_{21}}\right)i_1 + \left(L_{22}\frac{d^2}{dt^2} + R_{22}\frac{d}{dt} + \frac{1}{C_{22}}\right)i_2 = 0
\end{aligned} \tag{5.9}$$

We note that the matrix of Equations 5.8 and 5.9 is symmetric.

$$L_{12} = L_{21},\quad R_{12} = R_{21},\quad \frac{1}{C_{12}} = \frac{1}{C_{21}} \tag{5.10}$$

The above example makes it clear how to formulate the differential equations of the general n-mesh network. The equations will have the same form as Equation 5.9, but there will be n of them, and n currents i(t) will appear as the unknown functions of time.

$$\begin{aligned}
&\left(L_{11}\frac{d^2}{dt^2} + R_{11}\frac{d}{dt} + \frac{1}{C_{11}}\right)i_1 + \ldots\ldots + \left(L_{1n}\frac{d^2}{dt^2} + R_{1n}\frac{d}{dt} + \frac{1}{C_{1n}}\right)i_n = 0\\
&\text{---}\\
&\left(L_{n1}\frac{d^2}{dt^2} + R_{n1}\frac{d}{dt} + \frac{1}{C_{n1}}\right)i_1 + \ldots\ldots + \left(L_{nn}\frac{d^2}{dt^2} + R_{nn}\frac{d}{dt} + \frac{1}{C_{nn}}\right)i_{nn} = 0
\end{aligned} \tag{5.11}$$

To simplify future work we agree to use elastance S instead of capacitance C

$$S = \frac{1}{C} \tag{5.12}$$

and we rename the loop currents as follows:

$$\begin{bmatrix} i_1 \\ i_2 \\ \vdots \\ i_n \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = U_1 \tag{5.13}$$

The necessity for the subscript 1 in U_1 will become clear below, when U_1 will be used as a submatrix. Substituting from Equations 5.12 and 5.13 into the Equations 5.11 and carrying out the indicated differentiations we find

$$\begin{gathered} \left(L_{11}\ddot{u}_1 + R_{11}\dot{u}_1 + S_{11}u_1\right) + \cdots\cdots + \left(L_{1n}\ddot{u}_n + R_{1n}\dot{u}_n + S_{1n}u_n\right) = 0 \\ \text{- - - - - - - - - - - - - - - - - - - -} \\ \left(L_{n1}\ddot{u}_1 + R_{n1}\dot{u}_1 + S_{n1}u_1\right) + \cdots\cdots + \left(L_{nn}\ddot{u}_n + R_{nn}\dot{u}_n + S_{nn}u_n\right) = 0 \end{gathered} \tag{5.14}$$

Equations 5.14 can next be regrouped and written in matrix form as follows:

$$\begin{bmatrix} L_{11} \ldots\ldots L_{1n} \\ \ldots\ldots\ldots\ldots \\ L_{n1} \ldots\ldots L_{nn} \end{bmatrix} \begin{bmatrix} \ddot{u}_1 \\ \vdots \\ \ddot{u}_n \end{bmatrix} + \begin{bmatrix} R_{11} \ldots\ldots R_{1n} \\ \ldots\ldots\ldots\ldots \\ R_{n1} \ldots\ldots R_{nn} \end{bmatrix} \begin{bmatrix} \dot{u}_1 \\ \vdots \\ \dot{u}_n \end{bmatrix} + \begin{bmatrix} S_{11} \ldots\ldots S_{1n} \\ \ldots\ldots\ldots\ldots \\ S_{n1} \ldots\ldots S_{nn} \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = O$$

$$L\ddot{U}_1 + R\dot{U}_1 + SU_1 = O \tag{5.15}$$

The close analogy between Equations 5.3 and 5.15 is now clear. L, R, and S are matrices of inductive, resistive and capacitive circuit elements respectively. Their diagonal elements consist of sums of actual inductances, resistances and elastances taken around each loop. Thus, L_{22} is the sum of inductances taken around the second loop, and similarly for R_{22} and S_{22}. The off diagonal elements of the circuit matrices are the *negative* sums (see Equation 5.8) of the actual inductances etc. common to pairs of loops. Thus L_{ij} is the negative sum of inductances common to loops i and j, and similarly for R_{ij} and S_{ij}.

The next step in the formulation of the circuit equations is to remove the second time derivative and to replace Equation 5.15 by a form analogous to

Equation 5.5. This step is accomplished with the help of a substitution analogous to Equation 5.4. We let

$$\dot{U}_1 = L^{-1}U_2 \tag{5.16}$$

where U_2 is a column vector of n new functions of time, defined in terms of the n loop currents by Equations 5.16 and 5.13. L^{-1} is the reciprocal of the inductance matrix of the circuit, assumed non-singular. Written out in greater detail Equation 5.16 becomes

$$\begin{bmatrix} \dot{u}_1 \\ \dot{u}_2 \\ \vdots \\ \dot{u}_n \end{bmatrix} = L^{-1} \begin{bmatrix} u_{n+1} \\ u_{n+2} \\ \vdots \\ u_{2n} \end{bmatrix}$$

The new functions of time u_{n+j} are labelled by subscripts from n + 1 to 2n, because we shall presently want to append them to the original set u_j as a continuation. We substitute from Equation 5.16 into Equation 5.15 and rearrange the latter as follows:

$$U_2 = -SU_1 - RL^{-1}U_2 \tag{5.17}$$

Using partitioned matrices this can be written in the form

$$\dot{U}_2 = -\begin{bmatrix} S & RL^{-1} \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \tag{5.18}$$

Finally both Equations 5.16 and 5.18 can be assembled into the single matrix equation

$$\begin{gathered} \begin{bmatrix} \dot{U}_1 \\ \dot{U}_2 \end{bmatrix} = \begin{bmatrix} O & L^{-1} \\ -S & -RL^{-1} \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \\ \dot{U} = AU \end{gathered} \tag{5.19}$$

It should now be clear that the column vectors U_1 and U_2 were given extra subscripts to bring out the partitioned structure of Equation 5.19. The column vector on the right side of Equation 5.19 now contains 2n time functions ranging from $u_1(t)$ to $u_{2n}(t)$. The first set of n functions U_1 consists of the original loop currents i(t), while the second set U_2 consists of their first derivatives $\frac{d}{dt}$ i(t). The above form of the differential equations of the n-mesh network is of the same symbolic form as Equation 5.5 for the single loop circuit.

From Equation 5.19 we see that the n-mesh network is a dynamical system having 2n *degrees of freedom*, or *dimensions*, since 2n functions, or dependent variables $u_j(t)$, appear in the column vector U. The network is said to be *linear* because the dependent variables appear in the first power only in Equation 5.19. We say that the network has *properties independent of time*, because all elements of the square matrix in Equation 5.19 are constant, as they must be, since they are made up of the individual fixed inductances, resistances and capacitances of the circuit.

It is instructive to consider Equation 5.19 from the point of view of linear transformations. The matrix A has the remarkable property of transforming the column vector of time functions U into the column vector of their first derivatives $\dot{U}$, without containing any differential operators. This fact becomes less amazing as soon as we write down Equation 5.19 for the simplest case of order 1×1. We then obtain the scalar differential equation

$$\dot{u} = au \tag{5.20}$$

The reader will recall that a relation of this form is satisfied by the exponential function e^{at}.

$$\dot{e}^{at} = ae^{at} \tag{5.21}$$

Equation 5.21 is a special case of Equation 5.19: the exponential function is transformed into its derivative without the use of a differential operator.

It is tempting to jump to the conclusion that Equation 5.19 might have a solution of the same symbolic form as Equation 5.21. In the following section we show that this conclusion is quite justifiable.

5.2 SOLUTION OF THE EQUATIONS OF THE FORCE FREE NETWORK

In this section the solution of Equation 5.19 for the general network will be obtained.

We use the method of solution of differential equations by infinite series. Assuming that all the functions of time, contained in the column vector $U(t)$, can be expanded in Taylor series about t = o we can write, by Equation 4.129,

$$U(t) = U(0) + \frac{t}{1!}\dot{U}(0) + \frac{t^2}{2!}\ddot{U}(0) + \cdots\cdots \tag{5.22}$$

Now, from Equation 5.19 we obtain

$$\dot{U} = AU$$

$$\ddot{U} = A\dot{U} = A^2U \text{ etc.}$$

In general the n-th derivative of $U(t)$ is

$$\frac{d^n}{dt^n} U(t) = A^n U(t) \tag{5.23}$$

Evaluating the time derivatives of $U(t)$ at t = o from Equation 5.23, and substituting them into Equation 5.22 we find

$$U(t) = (1 + \frac{1}{1!}At + \frac{1}{2!}A^2t^2 + \ldots\ldots) U(0)$$

$$U(t) = e^{At}\, U(0) \tag{5.24}$$

Equation 5.24 constitutes the general solution of the network Equation 5.19. provided the initial conditions are given in the form of the column vector $U(0)$. Recalling that this vector consists of the loop currents and their first derivatives, we see that values of these quantities at the time t = 0 must be known. Since the vector U has 2n elements, the number of constants required to fix the solution is 2n, reflecting the 2n-dimensional character of the problem.

When applying Equation 5.24 to the description of the network, it is helpful to give it a geometrical interpretation. The state of the network is fully determined at any instant of time by the 2n-dimensional vector $U(t)$. Equation 5.24 expresses this state vector as a linear transformation of the initial constant state vector $U(0)$. The exponential transformation matrix e^{At} determines the time variation of the network, once its initial state is known.

Instead of using the terminology of linear transformations it is often convenient to use the concept of linear operators. Thus e^{At} may be called a time dependent operator. It generates the state vector of the network at any instant, by operating on its given initial state at the time t = 0.

Taking the geometrical view of our problem it is possible to obtain the solution of Equation 5.19 in a form which is both simpler mathematically than Equation 5.24, and easier to interpret physically. Towards the end of Section 4.8 it was explained that a linear transformation assumes a particularly simple form in a coordinate system defined by the eigenvectors of the transformation matrix. Let us, therefore, change the coordinates of our problem according to the law

$$V = Q^{-1}U \tag{5.25}$$

where Q is the matrix composed of eigenvectors of A. According to Equation 5.25 every one of the new co-ordinates v_i is a linear combination of some or all of the original coordinates u_j. Every new coordinate $v_i(t)$ can be said to contain an admixture from some or all of the loop currents and their derivatives $u_j(t)$. In a physical sense each co-ordinate v_i contains some information about the electrical state of every loop of the circuit. Hence, the new co-ordinates would appear to be rather complicated electrical quantities, but we shall see below that they can be given a useful interpretation. Taking the reciprocal of Equation 5.25.

$$U = QV \tag{5.26}$$

we see that every u_j can be expressed as a linear superposition of the v_i. This fact, together with a solution of Equation 5.19 in the changed co-ordinates, will provide a clue to the physical meaning of the v_i.

Substituting from Equation 5.26 into Equation 5.19 we find

$$Q\dot{V} = AQV \tag{5.27}$$

Premultiplying both sides by Q^{-1} we obtain the equation of the n-loop network in the new coordinates.

$$\dot{V} = Q^{-1}AQV$$
$$\dot{V} = \Lambda V \tag{5.28}$$

By virtue of the fact that Λ is diagonal, Equation 5.28 splits into 2n separate scalar equations for the new coordinates v_i.

$$\dot{v}_i = \lambda_i v_i, \quad (i = 1, 2, \ldots\ldots 2n) \tag{5.29}$$

Each of these equations can now be solved by elementary methods of integration. However, it is more instructive to treat them collectively and solve Equation 5.28 by the method that led to Equation 5.24. In other words we consider Equation 5.24 to be a general formula applicable to the Equation 5.28. The solution is

$$V(t) = e^{\Lambda t}\, V(0) \tag{5.30}$$

By Equation 4.120 $e^{\Lambda t}$ is a diagonal matrix having exponential functions of the form $e^{\lambda_i t}$ in its diagonal positions. Hence Equation 5.30 splits into 2n separate solutions for the coordinates v_i.

$$v_i(t) = e^{\lambda_i t} v_i(0), \quad (i = 1, 2, \ldots\ldots 2n) \tag{5.31}$$

The same solutions can be obtained by direct integration of Equation 5.29.

The final solution of the problem, the vector of loop currents and their first derivatives $U(t)$, can now be recovered through Equation 5.26.

$$U(t) = Q\, e^{\Lambda t}\, V(0) \tag{5.32}$$

Equation 5.32 can be rewritten in an alternative form using the fact that Q is composed of eigenvectors X_i of A. In the notation of Section 4.8 we have

$$U(t) = \sum_{i=1}^{2n} e^{\lambda_i t}\, v_i(0) X_i \tag{5.33}$$

As an example let us write out Equation 5.33 in full for the one loop network which has 2 independent coordinates.

$$\begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} = e^{\lambda_1 t} v_1(0) \begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix} + e^{\lambda_2 t} v_2(0) \begin{bmatrix} x_{12} \\ x_{22} \end{bmatrix} \tag{5.34}$$

Equation 5.33 displays *the time dependent vector of loop currents as a linear combination or superposition of the eigenvectors X_i of the circuit matrix A. The scalar multipliers, which go with each eigenvector, include the exponential time factor associated with the corresponding eigenvalue.*

Before discussing the physical interpretation of the Equations 5.31 and 5.33, several new terms must be introduced. The new coordinates v_i, which made the general solution of the network problem relatively simple, are called the *normal coordinates* of the problem. Each individual solution (Equation 5.31) is called a *normal mode* of the circuit. Every normal mode is characterised by an eigenvalue of the circuit matrix A. Equation 5.31 shows that the time dependence of every normal mode is decided by the eigenvalue to which it belongs, through the corresponding exponential time factor.

Since the exponential time factors are the most important part of the solution let us consider them more closely. The eigenvalues to which they belong are roots of an algebraic equation of degree 2n.

$$|A - \lambda I| = 0$$

$$\left\| \begin{matrix} -\lambda I & L^{-1} \\ -S & -RL^{-1} - \lambda I \end{matrix} \right\| = 0 \tag{5.35}$$

In general the roots are either real or complex, in which latter case they occur in complex conjugate pairs. For a perfectly general equation the real roots, and the real parts of complex roots, may be either positive or negative, but in the special case of the eigenvalue equation, (Equation 5.35), of a passive linear network, they are always negative. This statement can be proved from the form of the circuit matrix A, but here we accept it as already established. Writing the eigenvalues of the passive linear circuit in the form

$$\lambda_i = -\sigma_i + j\,\omega_i \tag{5.36}$$

where σ_i is assumed positive, the corresponding time dependent exponentials are

$$e^{-\sigma_i t}\, e^{j\omega_i t} \tag{5.37}$$

This is the familiar form of a sinusoidal oscillation of angular frequency ω_i, decaying exponentially with time according to the factor $e^{-\sigma_i t}$. The eigen-

values (Equation 5.36) are the *complex frequencies* of the circuit. Their imaginary parts ω_i are the natural oscillation frequencies of the network, while their real parts $-\sigma_i$ represent the damping due to circuit resistances.

The physical meaning of the normal modes is best seen on a special case. Let us assume that all the normal modes except one vanish. Referring to Equation 5.33 we find

$$\begin{aligned} U(t) &= e^{\lambda_k t}\, v_k(0)\, X_k \\ &= e^{-\sigma_k t}\, e^{j\omega_k t}\, v_k(0)\, X_k \end{aligned} \qquad (5.38)$$

where k is the label of the only non-vanishing normal mode. Concentrating our attention on the time dependent part of Equation 5.38, we see that all the loop currents, included in $U(t)$, oscillate at the frequency ω_k, and the oscillations decay at the rate $e^{-\sigma_i t}$ throughout the network.

In the general case each individual loop current in the network contains admixtures of all normal modes. Hence it oscillates in a complicated manner, the oscillations containing contributions at all the natural frequencies and decay rates of the circuit.

As an application of the method explained in this section the reader should write out in full the solution for the one-loop network of Fig. 5.1, as given by Equation 5.34. The eigenvalues of the circuit matrix should be computed from Equation 5.35, the eigenvectors of the same matrix should be found from the corresponding sets of homogeneous equations, and finally the scalar coefficients $v_i(0)$ should be determined from the initial state of the circuit.

Beyond this simplest case a detailed general solution of any network becomes too lengthy to attempt on paper. However, in specific cases electronic computers may be programmed to obtain numerical solutions.

5.3 MATRIX FORM OF THE LAPLACE TRANSFORMATION

It is assumed that the reader is familiar with the basic form of the Laplace transformation and its use in the analysis of linear networks. The object of this section is to introduce the Laplace transformation in matrix form, and to show how to apply it to the solution of the differential equations of passive linear circuits.

The basic relations of the Laplace transformation are

$$g(p) = \int_0^\infty e^{-pt} u(t)\, dt = \mathscr{L}(u) \qquad (5.39)$$

$$u(t) = \frac{1}{j2\pi} \oint e^{pt} g(p)\, dp = \mathscr{L}^{-1}(g) \qquad (5.40)$$

The contour over which the integration in Equation 5.41 is to be carried out contains all singularities of the integrand. The symbol $\mathscr{L}(u)$ stands for 'Laplace transform of u', where $u = u(t)$ is a time function. The symbol $\mathscr{L}^{-1}(g)$ stands for the 'inverse Laplace transform of g', where $g = g(p)$ represents the time function in the p-world, or the p-domain.

Given a column vector of time functions $U(t)$, their Laplace transforms $G(p)$ can be collected into a single matrix expression as follows:

$$G(p) = \int_0^\infty e^{-pt}\, I U(t)\, dt = \int_0^\infty e^{-pIt}\, U(t)\, dt = \mathscr{L}(U) \qquad (5.41)$$

$$U(t) = \frac{1}{j2\pi} \oint e^{pt}\, I G(p)\, dp = \frac{1}{j2\pi} \oint e^{pIt}\, G(p)\, dp = \mathscr{L}^{-1}(G) \qquad (5.42)$$

The reader should satisfy himself that the above expressions are meaningful and correct by referring back to Equations 4.115 and 4.120, and by writing out in full some examples of low order.

The transformations, represented by Equations 5.41 and 5.42, will now be applied to the solution of Equation 5.19. After this preliminary practice on the example of the differential equations of the force free network, we shall proceed to the problem of the circuit with forces in the following section.

Equation 5.19, rewritten here for convenient reference, is

$$\dot{U}(t) = A\, U(t) \qquad (5.43)$$

With the help of the Laplace transformation (Equation 5.41), both sides of Equation 5.43 are transformed into the p-domain. The transform of the right hand side is

$$\int_0^\infty e^{-pt}\, I A U(t)\, dt = A \int_0^\infty e^{-pt}\, I U(t)\, dt = A G(p) \qquad (5.44)$$

The circuit matrix A can be taken outside the integral sign because it commutes with the scalar matrix $e^{-pt}I$. The left hand side of Equation 5.43 is treated as follows:

$$\begin{aligned} \mathscr{L}(\dot{U}) &= \int_0^\infty e^{-pt}\, I \dot{U}(t)\, dt \\ &= \Big| e^{-pt}\, I U(t) \Big|_0^\infty + p \int_0^\infty e^{-pt}\, I U(t)\, dt \\ \mathscr{L}(\dot{U}) &= -U(0) + p\, G(p) \end{aligned} \qquad (5.45)$$

Integration by parts has been applied between the first and second lines above. The initial conditions of the problem now appear in the form of the definite integral $U(0)$.

Equating the transforms (Equations 5.44 and 5.45), we obtain

$$\mathrm{p}G(\mathrm{p}) - U(0) = A\,G(\mathrm{p})$$

Regrouping the terms so that $G(\mathrm{p})$ appears as a column vector of unknowns we find

$$(\mathrm{p}I - A)\,G(\mathrm{p}) = U(0) \tag{5.46}$$

Equation 5.46 is a system of 2n linear equations in 2n unknowns $G(\mathrm{p})$. Assuming that the matrix of the system, $(\mathrm{p}I - A)$, is nonsingular, Equation 5.46 can be solved.

$$G(\mathrm{p}) = (\mathrm{p}I - A)^{-1}U(0) \tag{5.47}$$

The solution of the differential equation (Equation 5.43) can now be obtained by transforming $G(\mathrm{p})$, as given by Equation 5.47, back into the time domain.

$$U(\mathrm{t}) = \frac{1}{\mathrm{j}2\pi}\oint \mathrm{e}^{\mathrm{pt}}\,I(\mathrm{p}I - A)^{-1}\,U(0)\,\mathrm{dp} \tag{5.48}$$

The above integral is easily evaluated after the matrix $(\mathrm{p}I - A)^{-1}$, considered as a function of A, has been decomposed into factors according to Equation 4.126.

$$\left(\mathrm{p}I - A\right)^{-1} = Q\left(\mathrm{p}I - \Lambda\right)^{-1}Q^{-1} \tag{5.49}$$

Here Λ is the diagonal form consisting of eigenvalues of the circuit matrix A. Q consists of the eigenvectors of A. Substituting from Equation 5.49 into Equation 5.48 we obtain

$$U(\mathrm{t}) = \frac{1}{\mathrm{j}2\pi}\oint \mathrm{e}^{\mathrm{pt}} I Q\left(\mathrm{p}I - \Lambda\right)^{-1}Q^{-1}\;U(0)\,\mathrm{dp} \tag{5.50}$$

Since Q is a constant matrix, it can be taken outside the integral sign by virtue of its commutability with the scalar matrix $\mathrm{e}^{\mathrm{pt}}I$. Similarly $Q^{-1}\,U(0)$ is unaffected by the integration, and since it commutes with the scalar factor dp it also can be shifted outside. Hence Equation 5.50 assumes the form

$$U(\mathrm{t}) = \frac{1}{\mathrm{j}2\pi} Q \oint \mathrm{e}^{\mathrm{pt}} I\left(\mathrm{p}I - \Lambda\right)^{-1}\mathrm{dp}\;Q^{-1}\;U(0) \tag{5.51}$$

Concentrating on the integrand in Equation 5.51 we observe that the reciprocal of the diagonal matrix $\mathrm{p}I - \Lambda$ is again diagonal.

$$(\mathrm{p}I - \Lambda)^{-1} = \begin{bmatrix} \mathrm{p} - \lambda_1 & 0 \cdots\cdots \\ 0 & \mathrm{p} - \lambda_2 \cdots\cdots \\ \cdots\cdots & \cdots\cdots \end{bmatrix}^{-1} = \begin{bmatrix} \dfrac{1}{\mathrm{p} - \lambda_1} & 0 \cdots\cdots \\ 0 & \dfrac{1}{\mathrm{p} - \lambda_2} \cdots\cdots \\ \cdots\cdots & \cdots\cdots \end{bmatrix} \tag{5.52}$$

The integral in Equation 5.51 therefore has the form

$$\oint \mathrm{e}^{\mathrm{pt}} I\left(\mathrm{p}I - \Lambda\right)^{-1}\mathrm{dp} = \oint \begin{bmatrix} \dfrac{\mathrm{e}^{\mathrm{pt}}}{\mathrm{p} - \lambda_1} & 0 \cdots\cdots \\ 0 & \dfrac{\mathrm{e}^{\mathrm{pt}}}{\mathrm{p} - \lambda_2} \cdots\cdots \\ \cdots\cdots & \cdots\cdots \end{bmatrix} \mathrm{dp} \tag{5.53}$$

Following the rules of Section 4.10., the foregoing integral can be evaluated element by element, using the method of residues. From the theory of functions of a complex variable we recall that the residue of a term like

$$\frac{\mathrm{e}^{\mathrm{pt}}}{\mathrm{p} - \lambda_\mathrm{i}}$$

at the singularity $\mathrm{p} = \lambda_\mathrm{i}$, is $\mathrm{e}^{\lambda_\mathrm{i}\mathrm{t}}$. Hence the integral in Equation 5.53 is

$$\oint \mathrm{e}^{\mathrm{pt}} I\left(\mathrm{p}I - \Lambda\right)^{-1}\mathrm{dp} = \mathrm{j}2\pi \begin{bmatrix} \mathrm{e}^{\lambda_1\mathrm{t}} & 0 \cdots\cdots \\ 0 & \mathrm{e}^{\lambda_2\mathrm{t}} \cdots\cdots \\ \cdots\cdots & \cdots\cdots \end{bmatrix} = \mathrm{j}2\pi\mathrm{e}^{\Lambda\mathrm{t}} \tag{5.54}$$

Substitution from Equation 5.54 into Equation 5.51 yields the result

$$U(\mathrm{t}) = Q\mathrm{e}^{\Lambda\mathrm{t}}\,Q^{-1}\,U(0) \tag{5.55}$$

whence, by Equation 4.126, we obtain the final solution of the system of differential equations (Equation 5.43).

$$U(\mathrm{t}) = \mathrm{e}^{A\mathrm{t}}\;U(0) \tag{5.56}$$

This agrees with the solution Equation 5.24, found in the preceding section by the method of power series.

The reader may observe that the decomposition of the matrix $(pI-\Lambda)^{-1}$, applied above, is equivalent to a change of coordinates. Rewriting Equation 5.55 in the form

$$Q^{-1}U(t) = e^{\Lambda t}Q^{-1}U(0) \tag{5.57}$$

we obtain the solution of Equation 5.43 in terms of the normal coordinates introduced by Equation 5.25.

$$V(t) = e^{\Lambda t}\,V(0)$$

This is the same as Equation 5.30.

Before going on to the subject of circuits with forces, we mention a more concise notation that may be used to write the transforms Equation 5.41 and 5.42. Instead of writing the scalar variable p separately, we define the scalar matrix

$$P = pI$$

The Equations 5.41 and 5.42 then assume the symbolic form

$$G(P) = \int_0^\infty e^{-Pt}\,U(t)\,dt$$

$$U(t) = \frac{1}{j2\pi}\oint e^{Pt}\,G(P)\,dP$$

It is found that the above relations, including the composite variable P, can be handled by the usual rules of integration, without regard to the matrix character of the integrands. The simplification achieved by following this symbolic method is sometimes worth while, but in an introductory treatment it is preferable to use the longer forms represented by Equations 5.41 and 5.42. They bring out fully the matrix character of all manipulations, which it is essential to practice in the initial stages.

5.4 DIFFERENTIAL EQUATIONS OF THE NETWORK WITH FORCES AND THEIR SOLUTION

In Section 5.1 we set up the differential equations of linear passive circuits without any electromotive forces present. In this section we extend our equations to include forces applied to the network. We assume that the applied forces are functions of time, including the possibility of d.c. e.m.fs.

We apply Kirchoff's Voltage Law to the general n-mesh network, and obtain its equations through the procedure that led to Equation 5.15. Writing $U_1(t)$ for the column vector of loop currents, the equations of the circuit with forces are

$$L\dot{U}_1(t) + RU_1(t) + S\int U_1(t)\,dt = E(t) \tag{5.58}$$

where $E(t)$ is the column vector of e.m.fs present in the individual loops. Differentiating both sides of Equation 5.58 with respect to time and omitting the functional notation we obtain

$$L\ddot{U}_1 + R\dot{U}_1 + SU_1 = \dot{E} \tag{5.59}$$

Making now the substitutions

$$\begin{aligned}\dot{U}_1 &= L^{-1}U_2\\ \ddot{U}_1 &= L^{-1}\dot{U}_2\end{aligned} \tag{5.60}$$

we can write Equation 5.59 in the form

$$\dot{U}_2 = -SU_1 - RL^{-1}U_2 + \dot{E}$$

or

$$\dot{U}_2 = -\begin{bmatrix} S & RL^{-1}\end{bmatrix}\begin{bmatrix} U_1\\ U_2\end{bmatrix} + \dot{E} \tag{5.61}$$

Equations 5.60 and 5.61 can be collected in the single matrix equation

$$\begin{bmatrix}\dot{U}_1\\ \dot{U}_2\end{bmatrix} = \begin{bmatrix} 0 & L^{-1}\\ -S & -RL^{-1}\end{bmatrix}\begin{bmatrix} U_1\\ U_2\end{bmatrix} + \begin{bmatrix} 0\\ \dot{E}\end{bmatrix}$$

$$\dot{U} = AU + \dot{E} \tag{5.62}$$

We must remember that in the last equation the symbol $\dot{E}$ is a column vector of 2n elements, having zeros for the first n elements. The reader will find that there is no risk of confusion between this new column vector of 2n elements, and the original vector of n elements on the right side of Equation 5.59.

Comparing Equation 5.62 and Equation 5.19, we note that the difference is the presence of the additive term $\dot{E}$, representing the forces applied to the network. The difference is sufficiently important to warrant a verbal distinction. The system of *force free differential equations* represented by Equation 5.19, is said to be *homogeneous*, while the *system with forces* is said to be *inhomogeneous*. The importance of the distinction lies in the fact that Equations 5.62 cannot be solved by the method applied to the homogeneous problem in Section 5.2. However, they can be solved, at least in theory, with the help of the Laplace transformation.

To solve the inhomogeneous problem with the help of the Laplace transformation we follow the same method as was applied to the homogeneous equation in the preceeding section. Both sides of Equation 5.62 are first transformed into the p-domain. As a result a system of (non-differential) linear equations is obtained, with the transforms $G(p)$ of the required time functions $U(t)$ in the role of unknowns. The equations are solved by the algebraic methods of Chapter 4, thus expressing $G(p)$ in terms of known quantities. Finally the known form of $G(p)$ is transformed back into the time domain, yielding the desired time functions $U(t)$.

We start by transforming both sides of Equation 5.62 into the p-world. By Equation 5.44 we have

$$\mathscr{L}(A\ U) = A\mathscr{L}(U) = A\ G(p) \tag{5.63}$$

The time derivatives $\dot{U}$ and $\dot{E}$ are transformed into the p-world according to Equation 5.45.

$$\mathscr{L}(\dot{U}) = pG(p) - U(0) \tag{5.64}$$

$$\mathscr{L}(\dot{E}) = pH(p) - E(0) \tag{5.65}$$

In Equation 5.65 the column vector $H(p)$ represents the Laplace transform of the given (undifferentiated) e.m.fs.

$$H(p) = \mathscr{L}(E) = \int_0^\infty e^{-pt}\, IE(t)\,dt \tag{5.66}$$

Assembling the transforms represented by Equations 5.63, 5.64 5.65, according to Equation 5.62 we obtain the following system of linear equations

$$pG(p) - U(0) = AG(p) + pH(p) - E(0). \tag{5.67}$$

Regrouping the terms so as to put the equation into the standard form we find

$$(pI - A)G(p) = pH(p) - E(0) + .U(0)) \tag{5.68}$$

The solution of Equation 5.68 for the unknowns $G(p)$ is

$$G(p) = (pI - A)^{-1}(pH(p) - E(0) + U(0)) \tag{5.69}$$

Comparing Equation 5.69 with Equation 5.47 for the force free circuit it is seen that the additional feature is the presence of transforms of the applied e.m.f.s on the right hand side.

The final step in the solution of the inhomogenous problem is to transform Equation 5.69 back into the time domain.

$$U(t) = \mathscr{L}^{-1}(G) = \frac{1}{j2\pi}\oint e^{pt} I(pI - A)^{-1}(pH(p) - E(0) + U(0))dp \tag{5.70}$$

The integral in Equation 5.70 splits into three separate terms, the last two terms including the column vectors $E(0)$ and $U(0)$. The latter are of the same form as Equation 5.48 and can be evaluated according to the procedure following that equation. The results are

$$\frac{1}{j2\pi}\oint e^{pt} I(pI - A)^{-1}U(0)dp = e^{At}U(0) \tag{5.71}$$

$$\frac{1}{j2\pi}\oint e^{pt} I(pI - A)^{-1}E(0)dp = e^{At}E(0) \tag{5.72}$$

Substituting from Equations 5.71 and 5.72 into Equation 5.70 we obtain

$$U(t) = \frac{1}{j2\pi}\oint e^{pt}\, I(pI - A)^{-1}pH(p)dp + e^{At}\big(U(0) - E(0)\big) \tag{5.73}$$

This is as far as the inhomogeneous problem can be solved in general terms. To make progress beyond this point it is necessary to assume a specific form for the loop e.m.f.s $E(t)$, and to work out their transforms $H(p)$. After this is done the remaining integral in Equation 5.73 can be evaluated.

Before considering any special cases a word of general comment on Equation 5.73 can be passed. The last term of Equation 5.73 has the same time dependence as the solution of the force free network obtained and discussed in Section 5.2. There we found that the solution is transient in nature, the loop currents dying away according to decaying exponentials. In the presence of persisting forces we expect the currents to contain terms which do not decay with time. From the form of Equation 5.73, we may conclude that such steady state parts of the loop currents must be contained within the integral term, and that they will appear once the latter is evaluated.

As an example of the possibilities pointed out above we shall solve the problem for the familiar case of harmonic e.m.f.s, all of the same angular frequency ω. As the steady state part of the solution we expect to obtain the basic a.c. equations of the circuit, in the form used in Chapter 3.

The column vector of harmonic e.m.f.s, applied to the network, is written in exponential form.

$$E(t) = e^{j\omega t}\, E \tag{5.74}$$

On the right hand side of Equation 5.74 the vector E is independent of time. Its elements are the complex amplitudes of the individual e.m.f.s. The Laplace transform of Equation 5.74 is

$$H(p) = \int_0^\infty e^{-pt}\, I\ e^{j\omega t}\, E dt$$

$$= \int_0^\infty e^{(j\omega - p)t}\, I dt\ E$$

$$H(p) = \frac{1}{p - j\omega}E \tag{5.75}$$

To find the steady state of the harmonically driven network, the transform Equation 5.75 is substituted into the integral (Equation 5.73), and the latter evaluated by the method used in the preceding section. The decisive step in that solution was the decomposition of the matrix $(pI - A)^{-1}$ into factors including the diagonal form Λ of A. With the help of Equation 4.126, Integral 5.73 can be written

$$\frac{1}{j2\pi}\oint e^{pt}\, I(pI-A)^{-1}\frac{p}{p-j\omega}\, E\,dp =$$

$$= \frac{1}{j2\pi}\oint e^{pt}\, I\, Q^{-1}(pI-\Lambda)^{-1}Q\,\frac{p}{p-j\omega}\, E\,dp$$

$$= Q^{-1}\frac{1}{j2\pi}\oint \frac{pe^{pt}}{p-j\omega}(pI-\Lambda)^{-1}dp\; QE \tag{5.76}$$

Writing out the integral above in greater detail we find that it is of the form

$$\frac{1}{j2\pi}\oint\begin{bmatrix} \dfrac{pe^{pt}}{(p-j\omega)(p-\lambda_1)} & 0 \ldots\ldots\ldots\ldots \\ 0 & \dfrac{pe^{pt}}{(p-j\omega)(p-\lambda_2)} \ldots\ldots \\ & \end{bmatrix}dp$$

or more briefly

$$\frac{1}{j2\pi}\oint\left[\frac{pe^{pt}}{(p-j\omega)(p-\lambda_i)}\right]dp \tag{5.77}$$

The integrand (5.77) has two singularities at the points $p = j\omega$ and $p = \lambda_i$. Its residues at these singularities are respectively

$$\frac{j\omega e^{j\omega t}}{j\omega-\lambda_i} \quad\text{and}\quad \frac{\lambda_i e^{\lambda_i t}}{\lambda_i - j\omega}$$

Hence the integral (5.77) is the sum of two diagonal matrices.

$$\begin{aligned}\frac{1}{j2\pi}\oint\left[\frac{pe^{pt}}{(p-j\omega)(p-\lambda_i)}\right]dp &= \left[\frac{j\omega e^{i\omega t}}{j\omega-\lambda_i}\right] + \left[\frac{\lambda_i e^{\lambda_i t}}{\lambda_i - j\omega}\right] \\ &= j\omega e^{j\omega t} I\left(j\omega I - \Lambda\right)^{-1} - \left[\lambda_i e^{\lambda_i t}\right]\left(j\omega I - \Lambda\right)^{-1} \\ &= \left(j\omega e^{j\omega t} I - \left[\lambda_i e^{\lambda_i t}\right]\right)\left(j\omega I - \Lambda\right)^{-1} \\ &= \left(j\omega e^{j\omega t} I - \Lambda e^{\Lambda t}\right)\left(j\omega I - \Lambda\right)^{-1}\end{aligned} \tag{5.78}$$

Before substituting from Equation 5.78 into Equation 5.76 we observe that the term $\Lambda e^{\Lambda t}$ has the same time dependence as the solution of the passive circuit, Equation 5.32. As such it is transient in nature, and since we are trying to find the steady state solution of the driven circuit, we omit it. Writing $U^{(S)}(t)$ for the column vector of steady state currents and their derivatives, we obtain from Equation 5.76

$$U^{(S)}(t) = Q^{-1} j\omega e^{j\omega t} I\left(j\omega I - \Lambda\right)^{-1} QE$$

$$U^{(S)}(t) = j\omega e^{j\omega t} I\left(j\omega I - A\right)^{-1} E \tag{5.79}$$

The above is the steady state solution of the circuit in a rather unfamiliar form. To bring it into the usual form of a.c. circuit equations we proceed as follows. First we premultiply both sides of Equation 5.79 by $(j\omega I - A)$.

$$(j\omega I - A)\, U^{(S)} = j\omega e^{j\omega t} E$$

Next we recall the detailed form of the circuit matrix A and the column vector E from Equation 5.62 above, and write in partitioned form

$$\begin{bmatrix} j\omega I & -L^{-1} \\ S & j\omega I + RL^{-1}\end{bmatrix}\begin{bmatrix} U_1^{(S)} \\ U_2^{(S)}\end{bmatrix} = j\omega e^{j\omega t}\begin{bmatrix} O \\ E\end{bmatrix} \tag{5.80}$$

This is equivalent to two separate equations.

$$j\omega U_1^{(S)} - L^{-1} U_2^{(S)} = O$$

$$S U_1^{(S)} + \left(j\omega I + RL^{-1}\right) U_2^{(S)} = j\omega e^{j\omega t} E$$

Solving the first equation for the column vector $U_2^{(S)}$, substituting into the second, and dividing throughout by $j\omega$, we obtain

$$\left(j\omega L + R + \frac{1}{j\omega} S\right) U_1^{(S)}(t) = e^{j\omega t} E \tag{5.81}$$

The column vector of steady state loop currents $U_1^{(S)}(t)$ must vary with time according to the exponential form $e^{j\omega t}$. Extracting the exponential factor outside, the currents can be written in terms of complex amplitudes as follows

$$U_1^{(S)}(t) = e^{j\omega t} U_1^{(S)} = e^{j\omega t} I \tag{5.82}$$

where *I is the column vector of a.c. loop currents and not the unit matrix.* The sum of the circuit matrices in brackets can be written as a single impedance matrix.

$$j\omega L + R + \frac{1}{j\omega} S = Z \tag{5.83}$$

Substituting the newly defined symbols from Equations 5.82 and 5.83 into Equation 5.81, we arrive at the final result

$$Z\,I = E \tag{5.84}$$

This is the usual form of a.c. circuit equations introduced in matrix form in Chapter 3. It should be noted that the order of the system of linear equations (represented by Equation 5.84) is $n \times n$, equal to the number of loops in the network.

The foregoing example illustrates the Laplace transformation method of solving the inhomogeneous system of differential equations, which arises when a circuit is driven by e.m.f.s given as functions of time. In cases in which the forces are not harmonic the integration problem may be more difficult, but the procedure remains the same in principle.

Chapter 6

Wave Matrices

The two-port networks discussed in Chapter 3 are the simplest case of networks or junctions which can have any number of ports. The theory of such multiport networks finds its most important application to junctions of transmission lines and waveguides. Since the terminal conditions in waveguide junctions are more conveniently described in terms of wave amplitudes than voltages and currents, it is to be expected that a new type of matrix will be required to relate them. The elements of the new matrices are not impedances, or current or voltage ratios, but coefficients relating wave amplitudes at the ports of networks, joining wave propagating lines. It is for this reason that the new matrices are called wave matrices.

The present chapter relies only to a small extent on the algebraic background of Chapter 4., it can, therefore, be read immediately after Chapter 3. The occasional gap in the mathematical argument should not deter a reader interested in such subjects as scattering matrices.

6.1 MULTIPORT NETWORKS

In Chapter 3 the parameters of two-port networks were discussed at length. In Section 3.2 it was shown how a two-port network is formed by isolating two loops of a general n-mesh circuit, and treating them as the input and output of a 'black box'. The procedure outlined there can be extended to form 'black boxes' with more than two ports. As an example of the method the equations of a three-port network will now be derived.

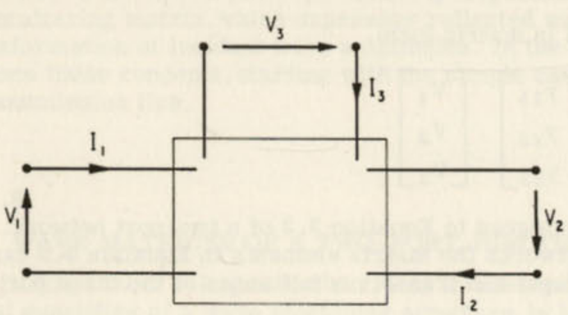


Fig. 6.1

The 'black box' of Fig. 6.1 is assumed to be a general passive linear circuit, with three of its loops taken outside to provide access ports. Assuming further that the three voltages marked in Fig. 6.1 represent the only generators applied to the network, its equations are

$$\begin{bmatrix} Z_{11} & Z_{12} & \cdots & Z_{1n} \\ Z_{21} & Z_{22} & \cdots & Z_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ Z_{n1} & Z_{n2} & \cdots & Z_{nn} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ \vdots \\ I_n \end{bmatrix} = \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad (6.1)$$

according to Equation 3.10. As with two-port networks we are interested in the currents flowing in the three terminal ports only. Hence we solve Equation 6.1 for the three currents I_1, I_2 and I_3.

$$I_1 = \frac{|Z_{11}|}{|Z|}V_1 + \frac{|Z_{21}|}{|Z|}V_2 + \frac{|Z_{31}|}{|Z|}V_3$$

$$I_2 = \frac{|Z_{12}|}{|Z|}V_1 + \frac{|Z_{22}|}{|Z|}V_2 + \frac{|Z_{32}|}{|Z|}V_3$$

$$I_3 = \frac{|Z_{13}|}{|Z|}V_1 + \frac{|Z_{23}|}{|Z|}V_2 + \frac{|Z_{33}|}{|Z|}V_3 \qquad (6.2)$$

The coefficients on the right hand side of Equations 6.2 have the dimensions of admittance. They are the admittance parameters of the three-port network of Fig. 6.1. Writing them in the form

$$y_{ij} = \frac{|Z_{ji}|}{|Z|},$$

Equations 6.2 can be put in matrix form.

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix} \qquad (6.3)$$

The above equation is analogous to Equation 3.3 of a two-port network. By analogy with two-port networks the matrix elements in Equation 6.3 can be called the short circuit input and transfer admittances of the three port network.

The 3×3 matrix of admittance parameters of a passive linear three-port network can be shown to be symmetric by the method used in Section 3.2 to establish the symmetry of the y-matrix of a two-port network.

Extending the methods developed in Chapter 3., we can go on and form matrices of other parameters of the three-port network. Thus Equation 6.3 can be solved for the terminal voltages, expressing the latter as linear transformations of terminal currents.

$$V = Y^{-1}I = ZI \qquad (6.4)$$

In this way the impedance parameters of the three-port network are formed. Continuing along this line various sets of mixed parameters can be obtained. The total number of these sets of three-port network parameters can be computed, if we note that it must be equal to the number of combinations of 3 objects, taken out of a group of 6, there being 6 terminal voltages and currents. The number is

$$\binom{6}{3} = \frac{4.5.6}{1.2.3} = 20$$

The foregoing example shows how a network with n ports can be treated. The mathematical procedure is the same, but the order of parameter matrices increases. Also the number of possible sets of parameters increases steeply with the number of access ports.

The theory of n-port networks finds its most important application to microwave junctions. However, instead of using voltages and currents, as the basic electrical quantities at the ports, it is preferable to use wave amplitudes. The latter are measurable in waveguides, whereas voltages and currents are not even uniquely defined.

The wave amplitudes at the ports of a junction are related by matrices, whose elements are neither impedances nor other parameters based on voltages and currents. Since their elements relate wave amplitudes, the new matrices are called *wave matrices*. A particularly important type of wave matrix is the scattering matrix, which expresses reflected wave amplitudes as a linear transformation of incident wave amplitudes. In the sections to follow we introduce these concepts, starting with the simple case of wave propagation on a transmission line.

6.2 WAVE MATRICES OF A TWO-PORT JUNCTION

The passage from voltages and currents to wave amplitudes, the basic electrical quantities of a wave sustaining structure, is best traced on the example of the transmission line. We recall from transmission line theory that the

voltage and current at the point x of a lossfree line, have the following exponential form:

$$V(x,t) = Ae^{(-j\beta x + j\omega t)} + Be^{(j\beta x + j\omega t)}$$

$$I(x,t) = \frac{1}{Z_0}\left(Ae^{(-j\beta x + j\omega t)} - Be^{(j\beta x + j\omega t)}\right) \qquad (6.5)$$

where *peak values*, not r.m.s. values are used, and

Z_0 = characteristic impedance of the line-real for a loss free line.

β = phase constant of the line;

ω = frequency of the wave carried by the line.

We further recall that the exponential term $e^{(-j\beta x + j\omega t)}$ represents a forward wave relative to the positive direction of x, while $e^{(j\beta x + j\omega t)}$ represents a reflected wave.

Equations 6.5 can be rewritten as follows:

$$V(x,t) = \sqrt{Z_0}\left(\frac{A}{\sqrt{Z_0}}e^{(-j\beta x + j\omega t)} + \frac{B}{\sqrt{Z_0}}e^{(j\beta x + j\omega t)}\right)$$

$$I(x,t) = \frac{1}{\sqrt{Z_0}}\left(\frac{A}{\sqrt{Z_0}}e^{(-j\beta x + j\omega t)} - \frac{B}{\sqrt{Z_0}}e^{(j\beta x + j\omega t)}\right) \qquad (6.6)$$

At this point we observe that there are only two distinct terms in the bracketed expressions of Equations 6.6, one representing a forward wave, the other a reflected wave. To economise on writing let us introduce new symbols.

$$\frac{A}{\sqrt{Z_0}}e^{(-j\beta x + j\omega t)} = a(x,t) = a$$

$$\frac{B}{\sqrt{Z_0}}e^{(j\beta x + j\omega t)} = b(x,t) = b \qquad (6.7)$$

The quantities a and b, defined by Equation 6.7, are called *normalised wave amplitudes*. The voltage and current equations (Equations 6.6) can now be written in terms of wave amplitudes.

$$V = \sqrt{Z_0}\,(a + b)$$

$$I = \frac{1}{\sqrt{Z_0}}(a - b) \qquad (6.8)$$

Conversely, solution of Equations 6.8, for the wave amplitudes yields the expressions

$$a = \frac{1}{2\sqrt{Z_0}}(V + Z_0 I)$$

$$b = \frac{1}{2\sqrt{Z_0}}(V - Z_0 I) \qquad (6.9)$$

Fig. 6.2 displays the wave amplitudes as electrical quantities of the transmission line. At the point x on the line, either the voltage and current, or the wave amplitudes define the electrical state of the line.

In the general case, in which both forward and reflected waves are present on a line, we can form the ratio of the reflected to forward wave amplitudes at any given point.

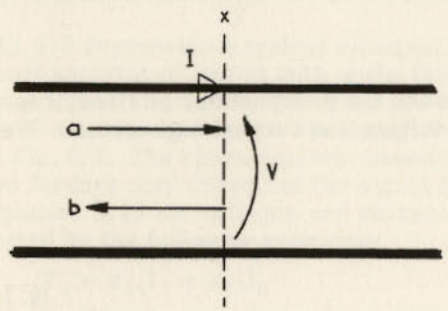


Fig. 6.2

Referring to Fig. 6.2 we can write, with the help of Equations 6.9

$$s = \frac{b(x,t)}{a(x,t)} = \frac{V(x,t) - Z_0 I(x,t)}{V(x,t) + Z_0 I(x,t)} = \frac{Z_x - Z_0}{Z_x + Z_0} \qquad (6.10)$$

Z_x is the impedance at the point x of the line. It is written for the ratio $V(x)/I(x)$, after the exponential time factor has been cancelled out of Equation 6.10.

The ratio s is called the *scattering coefficient* at the point x of the line. It is identical with the complex reflection coefficient, familiar from transmission line theory, and it is a function of position on the line, but not of time. Rewriting Equation 6.10 in the form

$$b = s\,a, \qquad (6.11)$$

we note that the scattering coefficient relates linearly the forward and reflected wave amplitudes.

O

The normalised wave amplitudes, as defined by Equations 6.7, provide a direct measure of the power carried by a wave. To see this clearly let us compute the power on a matched (or infinitely long) line, which propagates the forward wave only. In this case the voltage and current assume the simple form

$$V = \sqrt{Z_0}\, a, \quad I = \frac{1}{\sqrt{Z_0}} a \tag{6.12}$$

whence the power is

$$P = \frac{1}{2} V \bar{I} = \frac{1}{2} a \bar{a} = \frac{1}{2} |a|^2 \tag{6.13}$$

since Z_0 is a real quantity.

By a similar calculation the power carried by the reflected wave is $\frac{1}{2} b \bar{b} = \frac{1}{2} |b|^2$.

Before introducing the two-port network fed by transmission lines, it is desirable to redefine somewhat the voltage and current expressions. We re-write Equations 6.8 in the form

$$\frac{V}{\sqrt{Z_0}} = a + b = V^{(n)}$$

$$\sqrt{Z_0} I = a - b = I^{(n)} \tag{6.14}$$

The symbols $V^{(n)}$ and $I^{(n)}$ can be called normalised voltage and current, although dimensionally they are neither. They can be used to describe the electrical state of a line just as well as the proper voltage and current. In fact, equations expressed in terms of them are mathematically simpler. From a practical point of view it is immaterial how the voltage and current are defined, since measurements of these quantities are rarely made directly. Practically all measurements are concerned with wave amplitudes and power. Now, it is easily seen that the latter is expressed by the same equation, regardless of which form of voltage and current is used.

$$P = \frac{1}{2} V \bar{I} = \frac{1}{2} \frac{V}{\sqrt{Z_0}} \sqrt{Z_0}\, \bar{I} = \frac{1}{2} V^{(n)} \bar{I}^{(n)} \tag{6.15}$$

The above expression follows from the real nature of the characteristic impedance of a lossless line.

Having defined wave amplitudes as the basic electrical quantities of a transmission line, we pass on to the problem of a two-port network fed by lines. Our object is to describe the phenomena at the terminals of the network in terms of incident and reflected wave amplitudes, instead of voltages and currents as was done in Chapter 3.

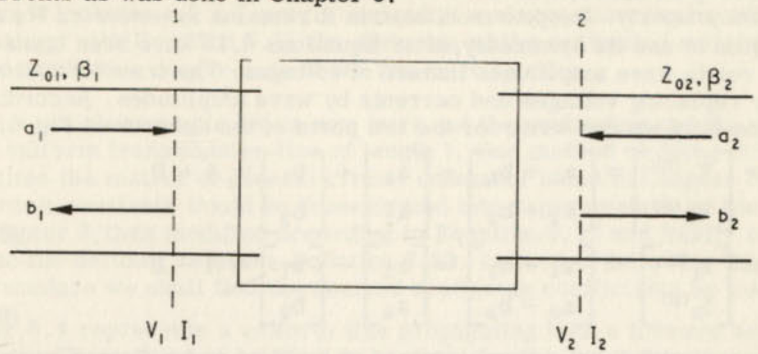


Fig. 6.3

Fig. 6.3 represents a typical situation. The network is assumed to include short sections of line at both ports, the actual terminals being marked by the dotted lines at 1 and 2. The lines are to be visualised as feeding the network, the *forward wave in each line being directed towards the network* as shown in Fig. 6.3. The characteristic impedances and propagation constants of the two feeders may differ, but the signal frequency is the same on both. By Equation 3.28 the voltages and currents at the ports of the network are connected by the following relations:

$$V_1 = z_{11} I_1 + z_{12} I_2$$

$$V_2 = z_{21} I_1 + z_{22} I_2 \tag{6.16}$$

Let us now modify these equations, by introducing the characteristic impedances of the lines.

$$\frac{V_1}{\sqrt{Z_{01}}} = \frac{z_{11}}{Z_{01}} \sqrt{Z_{01}} I_1 + \frac{z_{12}}{\sqrt{Z_{01} Z_{02}}} \sqrt{Z_{02}} I_2$$

$$\frac{V_2}{\sqrt{Z_{02}}} = \frac{z_{21}}{\sqrt{Z_{01} Z_{02}}} \sqrt{Z_{01}} I_1 + \frac{z_{22}}{Z_{02}} \sqrt{Z_{02}} I_2 \tag{6.17}$$

Using the newly defined normalised voltage and current, Equation 6.14, we can rewrite the above equations in simpler form.

$$V_1^{(n)} = \frac{z_{11}}{Z_{01}} I_1^{(n)} + \frac{z_{12}}{\sqrt{Z_{01} Z_{02}}} I_2^{(n)}$$

$$V_2^{(n)} = \frac{z_{21}}{\sqrt{Z_{01} Z_{02}}} I_1^{(n)} + \frac{z_{22}}{Z_{02}} I_2^{(n)}$$

$$V^{(n)} = Z I^{(n)} \tag{6.18}$$

The matrix of the above equation differs from the original impedance matrix of the two-port network, as used in Equations 6.16, but it shares with it an important property. Despite modification it remains *symmetric*. We shall have occasion to use its symmetry, after Equations 6.18 have been transformed to relate wave amplitudes instead of voltages. The transformation is effected by replacing voltages and currents by wave amplitudes. According to Equations 6.14 we can write for the two ports of the network of Fig. 6.3:

$$V^{(n)} = \begin{bmatrix} V_1{}^{(n)} \\ V_2{}^{(n)} \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = A + B \tag{6.19}$$

$$I^{(n)} = \begin{bmatrix} I_1{}^{(n)} \\ I_2{}^{(n)} \end{bmatrix} = \begin{bmatrix} a_1 - b_1 \\ a_2 - b_2 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = A - B \tag{6.20}$$

The matrices A and B are column vectors of incident and reflected wave amplitudes. They are to be used to describe the terminal conditions of the two port network instead of currents and voltages.

Substituting Equations 6.19 and 6.20 into Equation 6.18 we find

$$A + B = Z(A - B)$$

Next we rearrange the terms so that the column vector of reflected wave amplitudes B is on one side of the equation, while the column vector of incident wave amplitudes A is on the other side.

$$ZB + B = ZA - A$$

Using the unit matrix U the above equation can be rewritten in the form

$$(Z + U)B = (Z - U)A$$

Whence, on premultiplication of both sides by $(Z + U)^{-1}$, we obtain

$$B = (Z + U)^{-1}(Z - U)A \tag{6.21}$$

Equation 6.21 expresses the reflected wave amplitudes as a linear superposition of the incident wave amplitudes at the ports of the network. The matrix of Equation 6.21 is called the *scattering matrix* of the two-port network, and is usually denoted by the symbol S.

$$S = (Z + U)^{-1}(Z - U) \tag{6.22}$$

In terms of S, Equation 6.21 assumes the form

$$B = SA$$

$$\begin{aligned} b_1 &= s_{11}a_1 + s_{12}a_2 \\ b_2 &= s_{21}a_1 + s_{22}a_2 \end{aligned} \tag{6.23}$$

Equation 6.23 is an extension of Equation 6.11. It has the same linear algebraic form as Equation 6.11, but since it relates four quantities instead of two, it includes a 2×2 matrix instead of a single numerical coefficient. By analogy with Equation 6.11 the elements of the scattering matrix are called the *scattering coefficients* of the two-port junction.

As a simple example let us now work out the scattering matrix of a section of uniform transmission line of length l. One method of doing this would be to utilise the matrix of general circuit constants found in Chapter 3. The general circuit constants would be transformed into z-parameters by the methods of Chapter 3, then modified according to Equation 6.17 and finally substituted into the defining relation—Equation 6.22. Instead of following this lengthy procedure we shall find the desired scattering coefficients by inspection.

Fig. 6.4 represents a uniform line propagating both a forward and reflected wave. The section of length l, to be considered a two-port network, is marked by dotted lines. The wave amplitudes at the ports are labelled according to the convention defined in Fig. 6.3: the waves directed towards the ports are to be regarded as forward waves, designated by the symbol a_i regardless of whether they travel to the right or left. By Equations 6.6 and 6.7, but omitting the exponential time factor, we can write the amplitudes of the wave travelling to the right as follows:

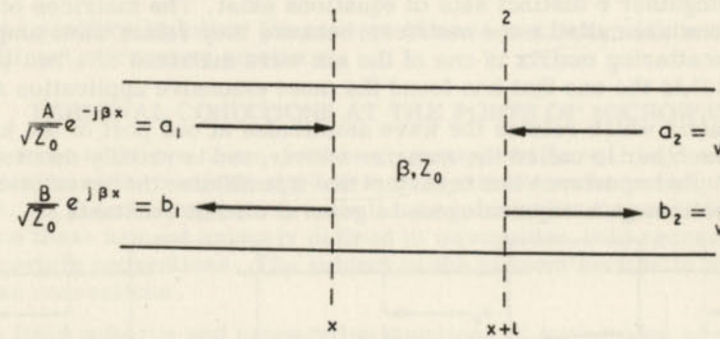


Fig. 6.4.

$$b_2 = \frac{A}{\sqrt{Z_0}}\, e^{-j\beta(x+l)} = a_1 e^{-j\beta l} + 0.\, a_2$$

Similarly for the wave travelling to the left we find

$$b_1 = \frac{B}{\sqrt{Z_0}} e^{j\beta x} = \frac{B}{\sqrt{Z_0}} e^{j\beta(x+l)}\ e^{-j\beta l} = a_2 e^{-j\beta l} + 0.\, a_1$$

Comparing the foregoing relations with Equations 6.23 we can write in matrix form

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 0 & e^{-j\beta l} \\ e^{-j\beta l} & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \qquad (6.24)$$

The matrix of Equation 6.24 is the scattering matrix of a section of line, shown in Fig. 6.4 as a two-port network. It is instructive to note in passing that it is symmetric.

The scattering matrix, derived above, is not the only type that can be used to relate wave amplitudes at the terminals of a two-port junction. Instead of expressing the reflected, or scattered, wave vector B in terms of the incident wave vector A, it is possible to express, say, the input waves a_1, b_1 in terms of the output waves a_2, b_2. The appropriate matrix can be obtained from the scattering coefficients by suitably rearranging Equations 6.23. Alternatively it can be worked out from the general circuit constants of the two-port network, by using them as the starting point instead of the z-parameters.

The four wave amplitudes define the electrical conditions at the terminals of a two-port network in the same way as the voltages and currents did in Chapter 3. Like the voltages and currents, any two of the wave amplitudes can be expressed linearly in terms of the remaining ones, by a continuation of the method suggested above. As in the case of voltages and currents, it is found that altogether 6 distinct sets of equations exist. The matrices of these equations are called *wave matrices*, because they relate wave amplitudes. The scattering matrix is one of the six wave matrices of a two-port junction, and it is the one that has found the most extensive application so far.

The wave matrix which relates the wave amplitudes at one port of the junction to those at the other is called the *transfer matrix*, and is usually denoted by the symbol T. Its importance lies in the fact that it facilitates the description of cascaded junctions in a way analogous to general circuit constants.

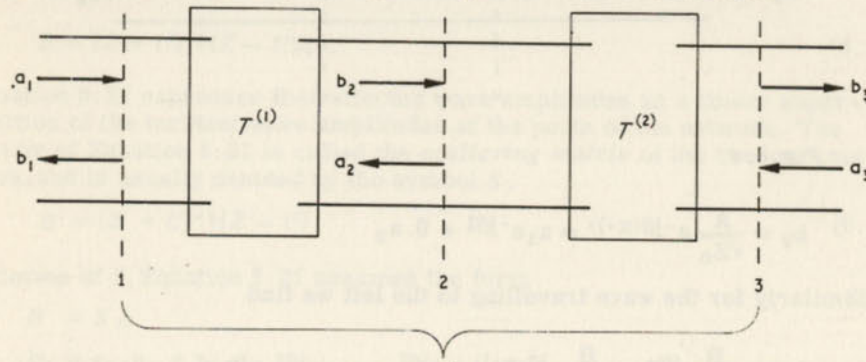


Fig. 6.5. $T = T^{(1)}\,T^{(2)}$

Referring to Fig. 6.5, we can write for each individual two-port:

$$\begin{bmatrix} a_1 \\ b_1 \end{bmatrix} = \begin{bmatrix} t_{11}^{(1)} & t_{12}^{(1)} \\ t_{21}^{(1)} & t_{22}^{(1)} \end{bmatrix} \begin{bmatrix} b_2 \\ a_2 \end{bmatrix}$$

$$\begin{bmatrix} b_2 \\ a_2 \end{bmatrix} = \begin{bmatrix} t_{11}^{(2)} & t_{12}^{(2)} \\ t_{21}^{(2)} & t_{22}^{(2)} \end{bmatrix} \begin{bmatrix} b_3 \\ a_3 \end{bmatrix} \qquad (6.25)$$

The reader should note that the amplitudes of waves travelling to the right of Fig. 6.5 are placed first in the column vectors of Equations 6.25, while those travelling to the left are placed second. As a result of this arrangement we can substitute the second equation into the first and obtain a matrix relating the input and output wave amplitudes of the cascade connection.

$$\begin{bmatrix} a_1 \\ b_1 \end{bmatrix} = T^{(1)} T^{(2)} \begin{bmatrix} b_3 \\ a_3 \end{bmatrix} = T \begin{bmatrix} b_3 \\ a_3 \end{bmatrix} \qquad (6.26)$$

The transfer matrix of the cascade connection is thus seen to be the product of the matrices of individual junctions, as shown in Fig. 6.5.

Other wave matrices of two-port junctions may find their application in special situations just as the mixed parameters of two-port networks have been applied to transistor circuits.

In the sections to follow the concept of wave matrices will be extended to multiport microwave junctions.

6.3 TERMINAL CONDITIONS AT THE PORTS OF MICROWAVE JUNCTIONS

The wave matrices of a n-port microwave junction will be derived by an extension of the methods applied to two-port networks in the preceding section. As the starting point we shall take terminal voltages and currents, and since these are not uniquely defined in waveguides, it is necessary to agree on certain conventions. The subject of the present section is to summarise these conventions.

The field patterns and propagation equations in waveguides are discussed at length in textbooks on microwave theory.* They are derived on the assumption that waveguides are loss free. For this reason the propagation equations of waveguide modes exhibit the same features as the equations of lossless transmission lines, used in the preceding section. Notably, the propagation constants are purely imaginary, and characteristic impedances are purely real. However, by contrast with transmission lines, which propagate the

---

* E.g. MONTGOMERY. C. G, DICKE. R. H, PURCELL. E. M, 'Principles of Microwave Circuits'; *M.I.T. Radiation Laboratory Series*, Vol. 8.

principal or TEM mode, there is no clearly defined voltage and current in waveguide modes. As a result of this ambiguity a variety of impedance and admittance matrices can be constructed for one and the same junction. Moreover, the matrices may not even by symmetric in the case of passive and linear junctions.

It can be shown that voltages, currents, and impedances, having the properties listed below, can be defined for waveguide modes.

1. The voltage and current of a unidirectional wave entering the i-th port of a junction yield the power flow through the expression

$$P_i = \frac{1}{2} V_i \bar{I}_i$$

where $\bar{I}_i$ is the complex conjugate of I_i, and peak values of the currents and voltages are used. This means that the power flow is given in terms of a wave amplitude by the formula

$$P_i = \frac{1}{2} a_i \bar{a}_i = \frac{1}{2} |a_i|^2$$

2. The characteristic impedance in a waveguide can always be set equal to unity
3. The impedance, and hence admittance matrices of passive linear junctions are symmetric

As a result of these assumptions the linear relations connecting currents and voltages at the terminals of a microwave junctions are of *the same form as Equations 6.3 and 6.4.* It should be remembered that the linear relations set up according to the above rules describe what happens at the ports of a microwave junction under conditions when more than one mode of propagation is present there. The currents and voltages are then defined for each individual mode, and all of them are included in their column vectors. In such cases the impedance and admittance matrices contain elements which connect individual modes, as well as separate ports, and their order is greater than the number of ports entering the junction.

On the basis of the rules outlined above it can be further shown that the voltage and current of a waveguide mode can be expressed in terms of its wave amplitudes by equations of the same form as Equations 6.14 of the transmission line. If the propagation mode under discussion is labelled by the subscript i, the equations are

$$V_i = a_i + b_i$$
$$I_i = a_i - b_i \qquad (6.27)$$

where a_i is the amplitude of the wave travelling towards the junction, and b_i is the amplitude of the wave reflected from the junction.

Equations 6.27 will provide the basis for the derivation of the scattering matrix of a n-port junction in the following section.

6.4 THE SCATTERING MATRIX OF A LINEAR AND RECIPROCAL MICROWAVE JUNCTION

Let us consider a microwave junction as shown diagrammatically in Fig. 6.6. It is assumed to be linear, passive and free of any nonreciprocal components such as ferrites. The junction has either n distinct ports, each carrying a single mode of propagation, or it has a smaller number of ports, some of which carry more than one mode. In the latter case each mode is treated as a separate port and labelled accordingly. This is possible by virtue of the orthogonality of waveguide modes. As a result of orthogonality individual modes cannot couple in waveguides, but they can interact within the junction proper. Besides coupling incident modes, microwave junctions are capable of exciting new modes, which are below the cut-off propagation properties of the input or output waveguides. Wherever such evanescent modes exist, it is assumed that the reference planes of the ports are sufficiently far inside the guides to exclude them.

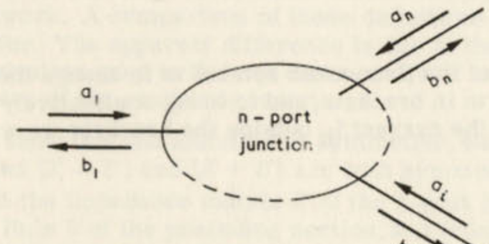


Fig. 6.6.

In what follows the scattering matrix of the n-port junction will be obtained, by basically the method applied to the transmission line in Section 6.2. The details of the algebraic argument will be varied, however, to illustrate the diversity of matrix methods.

As a first step in the derivation of the scattering matrix we solve Equations 6.27 of the preceding section for the incident and reflected wave amplitudes at the i-th port.

$$a_i = \frac{1}{2}(V_i + I_i)$$

$$b_i = \frac{1}{2}(V_i - I_i) \qquad (6.28)$$

Next we replace V_i by the corresponding row taken out of the impedance matrix, Equation 6.4, of the n-port junction. We find

$$\left.\begin{aligned} a_i &= \frac{1}{2}\left(\sum_{k=1}^{n} z_{ik}I_k + I_i\right) \\ b_i &= \frac{1}{2}\left(\sum_{k=1}^{n} z_{ik}I_k - I_i\right) \end{aligned}\right\} \qquad (6.29)$$

With the help of the Kronecker δ-symbol the above equations can be rewritten in the form

$$\left.\begin{aligned} a_i &= \frac{1}{2}\sum_{k=1}^{n}\left(z_{ik}I_k + \delta_{ik}I_k\right) \\ b_i &= \frac{1}{2}\sum_{k=1}^{n}\left(z_{ik}I_k - \delta_{ik}I_k\right) \end{aligned}\right\} \delta_{ik} = \begin{cases} 1 \text{ for } k = i \\ 0 \text{ for } k \neq i \end{cases}$$

We observe that the nett effect of the Kronecker symbol is to change the subscript of the second current term in brackets, and to bring it effectively under the summation symbol. Taking the current I_k outside the brackets we write

$$\left.\begin{aligned} a_i &= \frac{1}{2}\sum_{k=1}^{n}\left(z_{ik} + \delta_{ik}\right)I_k \\ b_i &= \frac{1}{2}\sum_{k=1}^{n}\left(z_{ik} - \delta_{ik}\right)I_k \end{aligned}\right\} \qquad (6.30)$$

The wave amplitudes a_i and b_i, considered as elements of column vectors A and B, are shown by Equations 6.30 to be the result of matrix products. In each case a square matrix of order $n \times n$ is postmultiplied by the column vector of currents I. In matrix symbols Equations 6.30 assume the form

$$\begin{aligned} A &= \frac{1}{2}(Z + U)I \\ B &= \frac{1}{2}(Z - U)I \end{aligned} \qquad (6.31)$$

where the unit matrix U replaces the Kronecker symbols.

Only one more step is required before the reflected or scattered waves b_i are expressed linearly in terms of the incident waves a_i. Solving the first of Equations 6.31 for the column vector of currents I we find

$$I = 2(Z + U)^{-1}A$$

Substitution of this expression into the second of Equations 6.31 yields the desired result

$$B = (Z - U)(Z + U)^{-1}A \qquad (6.32)$$

where

$$(Z - U)(Z + U)^{-1} = S \qquad (6.33)$$

is the *scattering matrix of the n-port junction.* Using this symbol Equation 6.32 is abbreviated to

$$B = S\,A \qquad (6.34)$$

Equation 6.34 applies to a general n-port junction, including the case $n = 2$, treated in Section 6.2, whence it follows that Equation 6.33 should have the same form as Equation 6.22, which defined the scattering matrix for a two-port network. A comparison of these definitions of S shows, however, that they differ. The apparent difference is due to the slightly different method of derivation used in the two cases, and will be explained in terms of the symmetry of the scattering matrix.

To show that the matrix S is symmetric, we observe at first that the matrices $(Z - U)$ and $(Z + U)$ are both symmetric. This follows from the fact that the impedance matrix Z of the n-port junction is assumed symmetric by Rule 3 of the preceding section, and from the fact that the addition of the diagonal matrix U does not affect symmetry. Moreover, by the theorem proved in Section 2.18 the reciprocal matrix $(Z + U)^{-1}$ is also symmetric. Bearing in mind these remarks we now take the transpose of S.

$$\begin{aligned} S^t &= \left\{ (Z - U)(Z + U)^{-1} \right\}^t \\ &= (Z + U)^{-1t}\,(Z - U)^t \\ S^t &= (Z + U)^{-1}(Z - U) \end{aligned} \qquad (6.35)$$

Next we postmultiply both sides of Equation 6.35 by the unit matrix, applying it to the right hand side in the form

$$U = (Z + U)(Z + U)^{-1}$$

Hence

$$S^t = (Z + U)^{-1}(Z - U)(Z + U)(Z + U)^{-1}$$

By direct expansion it is easily verified that the two factors in the middle of the above expression commute. Therefore we can write

$$S^t = (Z + U)^{-1}(Z + U)(Z - U)(Z + U)^{-1}$$

$$S^t = (Z - U)(Z + U)^{-1} = S \tag{6.36}$$

which establishes *the symmetry of the scattering matrix of a passive, reciprocal junction.*

The foregoing argument clarifies the difference between the form of the matrix S, obtained in this section and in Section 6.2. The slight variation in the method of derivation has led us to the transposed form of the scattering matrix.

The symmetry of the scattering matrix is a property of all linear and reciprocal junctions, including those that dissipate energy. In many applications it is useful to introduce a simplification by assuming that a junction is loss free. It is then found that the scattering matrix is unitary.

To prove this useful property of the S matrix we consider the power flow in and out of a junction. Writing the power entering the junction in terms of incident wave amplitudes we find

$$P_{in} = \frac{1}{2}\sum_i a_i\bar{a}_i \tag{6.37}$$

while the power leaving the junction is

$$P_{out} = \frac{1}{2}\sum_i b_i\bar{b}_i \tag{6.38}$$

The energy dissipated inside the junction is given by the difference of Equations 6.37 and 6.38.

$$P_{in} - P_{out} = \frac{1}{2}\sum_i (a_i\bar{a}_i - b_i\bar{b}_i) \tag{6.39}$$

In terms of matrix notation Equation 6.39 appears in the form

$$P_{in} - P_{out} = \frac{1}{2}(\bar{A}^tA - \bar{B}^tB) \tag{6.40}$$

where the power flow is expressed by scalar products (see p. 152) of the vectors of wave amplitudes. Substituting for the reflected wave amplitudes from Equation 6.34 we find

$$P_{in} - P_{out} = \frac{1}{2}(\bar{A}^tA - \bar{A}^t\bar{S}^tSA)$$

$$= \frac{1}{2}\bar{A}^t(U - \bar{S}^tS)A \tag{6.41}$$

The power leaving must equal the power entering a lossless junction, and the above difference must therefore vanish. Hence we obtain the condition

$$\bar{S}^t = S^{-1} \tag{6.42}$$

which proves the *unitary* property of the scattering matrix of a *dissipationless junction.*

The scattering coefficients derived above are directly related to the impedance parameters of a n-port junction. Like the latter they represent electrical properties of the junction and are, therefore, independent of such temporary conditions as the loads connected to any of the ports, or the method of feeding power into them. In general however, the scattering coefficients will be functions of frequency just as are the impedances.

Although the first and simplest scattering coefficient, introduced by Equation 6.10, happened to be identical with the complex reflection coefficient at a point of a transmission line, it must not be assumed that this is so in general. Only in special cases do some of the scattering coefficients happen to be reflection coefficients. Physically, the reason for this state of affairs is not difficult to see. The reflection coefficient at a port of a junction, will depend on the terminations or sources of energy connected to the remaining ports. As pointed out in the preceding paragraph the scattering coefficients cannot be so influenced, since they characterise the junction alone. To illustrate these observations further, we will discuss some examples.

Given an n-port junction as shown in Fig. 6.7, let us assume that energy is being fed into one port only, say the i-th port. The remaining ports are terminated in matched loads, so that only the reflected waves b_k are present there. In the i-th port, there is both an incident wave a_i and a reflected wave b_i. The scattering relation for the junction of Fig. 6.7 has the form

$$\begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ s_{21} & s_{22} & \cdots & s_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ s_{n1} & s_{n2} & \cdots & s_{nn} \end{bmatrix} \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ a_i \\ \cdot \\ 0 \end{bmatrix} \tag{6.43}$$

On multiplication this reduces to the following set of simple linear relations:

$$b_l = s_{li}\, a_i. \quad l = 1, 2, \ldots n \tag{6.44}$$

For $l = i$ we find

$$b_i = s_{ii}\, a_i \tag{6.45}$$

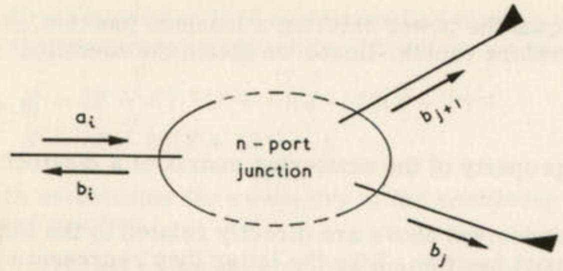


Fig. 6.7

In this case the diagonal element s_{ii} is identical with the reflection coefficient at the corresponding port.

$$\rho_i = \frac{b_i}{a_i} = s_{ii} \tag{6.46}$$

On the other hand let us assume that the termination at the output of one of the ports is not matched, so that both a reflected and a forward wave is present there. Giving this particular port the label, i + 1, we can write the scattering relation in the form

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ s_{21} & s_{22} & \cdots & s_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ s_{n1} & s_{n2} & \cdots & s_{nn} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ a_i \\ a_{i+1} \\ 0 \end{bmatrix} \tag{6.47}$$

Multiplication yields the equations

$$b_l = s_{li}\, a_i + s_{li+1}\, a_{i+1}, \quad l = 1, 2, \ \ldots . \ n \tag{6.48}$$

In this case it is not possible to form a simple relation of the form of Equation 6.45, connecting the incident and reflected wave amplitudes in the i-th or (i + 1)-st ports. Hence none of the scattering coefficients s_{ii} or $s_{i+1}a_{i+1}$ can be identified with the reflection coefficient in the corresponding port. We conclude, therefore, that a diagonal element of *S* becomes the reflection coefficient at the corresponding port only when all the remaining ports are matched.

An observation of practical importance regarding the diagonal elements of the scattering matrix can now be made. Referring to Fig. 6.7, we would not expect to find a reflected wave at the i-th port unless there are internal reflections in the junction. Conversely if we know that a given *junction has no internal reflections, we can immediately conclude that the diagonal elements of its S matrix must vanish.* This fact will be utilised below, when the scattering matrices of some specific junctions will be worked out.

To give a concrete example of the foregoing observations let us analyse a specific two-port junction. This may take the form of a waveguide to coaxial transition, an impedance transformer, or simply a waveguide bend which is not very well matched. Fig. 6.8 shows a waveguide bend. We assume that the output of port 2 is terminated in such a way that a wave of amplitude a_2 is reflected back towards the bend.

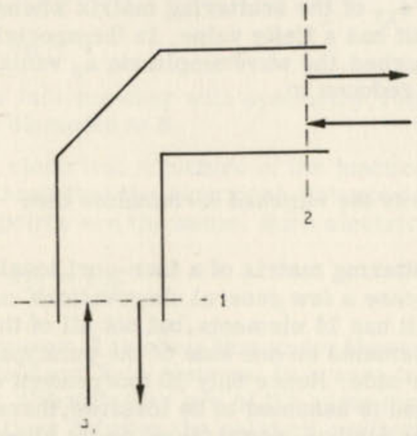


Fig. 6.8

The reflection coefficient at the output port, *looking towards it,* is

$$\rho_2 = \frac{b_2}{a_2} \tag{6.49}$$

It should be carefully noted that this quantity is the reciprocal of the reflection coefficient of the termination.

The scattering equations for the bend of Fig. 6.8 are

$$\begin{aligned} b_1 &= s_{11}a_1 + s_{12}a_2 \\ b_2 &= s_{21}a_1 + s_{22}a_2 \end{aligned} \tag{6.50}$$

We use the first of Equations 6.50 to form the reflection coefficient at the input of the two-port.

$$\rho_1 = \frac{b_1}{a_1} = s_{11} + s_{12}\frac{a_2}{a_1} \tag{6.51}$$

With the help of the second of Equations 6.50 and Equation 6.49 we eliminate the wave amplitudes a_1 and a_2 appearing on the right side of Equation 6.51. In the process the wave amplitude b_2 cancels out, and we arrive at the expression

$$\rho_1 = s_{11} + \frac{s_{12}^2}{\rho_2 - s_{22}} \tag{6.52}$$

where use has been made of the symmetry of S when putting $s_{21} = s_{12}$.

Equation 6.52 shows that the reflection coefficient at the input of the bend is not identical with the diagonal element s_{11} of the scattering matrix whenever the reflection coefficient ρ_2 at the output has a finite value. In the special case when the output termination is matched, the wave amplitude a_2 vanishes, ρ_2 becomes infinite, and Equation 6.52 reduces to

$$\rho_1 = s_{11}$$

The reflection coefficient looking *towards the matched termination* then equals zero.

As a final example we consider the scattering matrix of a four-port lossless junction. Before discussing a specific case a few general observations can be made. As the matrix is of order 4×4 it has 16 elements, but not all of them are independent. By symmetry the 6 elements on one side of the principal diagonal are equal to those on the other side. Hence only 10 independent elements remain. Furthermore the junction is assumed to be lossfree, therefore its matrix is unitary. This fact imposes further restrictions on its elements, which can be expressed by an application of Equation 4.108.

$$\sum_{i=1}^{4} \bar{s}_{im} s_{in} = \delta_{mn} \tag{6.53}$$

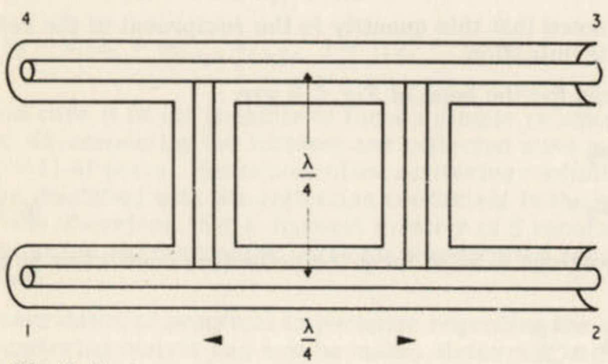


Fig. 6.9

As the above conditions are best used in a specific case we go on to formulate the scattering matrix of a 3db or half power directional coupler, sometimes called a hybrid junction. Here we consider perhaps the simplest form of such a junction, made up of sections of transmission line, as shown in Fig. 6.9. We assume that there are no internal reflections in the coupler, so that all the energy entering, say, port 1, appears at ports 2 and 3, and is equally split between them. From this assumption we can immediately deduce that each diagonal element of the S matrix must vanish, since it represents the reflection coefficient at the corresponding port, with the remaining ports matched (see p. 197).

$$s_{ii} = 0 \tag{6.54}$$

This fact, together with symmetry, reduces the number of independent matrix elements to 6.

The electrical structure of the junction provides further information, if it is assumed that the electrical distances between ports and the nearest branching points are the same. Such electrical symmetry entails the equality

$$s_{21} = s_{34} \tag{6.55}$$

Moreover, it is clear that under these conditions the waves appearing at ports 2 and 3, in response to a wave incident on ports 1 or 4, are in quadrature, although they are of the same magnitude. Hence we obtain the following relations between the relevant scattering coefficients:

$$s_{31} = e^{-j\pi/4}\, s_{21}$$

$$s_{24} = e^{-j\pi/4}\, s_{34} = e^{-j\pi/4}\, s_{21} \tag{6.56}$$

The scattering matrix of the coupler of Fig. 6.9 can now be written in the form:

$$\begin{bmatrix} 0 & s_{21} & e^{-j\pi/4}s_{21} & 0 \\ s_{21} & 0 & 0 & e^{-j\pi/4}s_{21} \\ e^{-j\pi/4}s_{21} & 0 & 0 & s_{21} \\ 0 & e^{-j\pi/4}s_{21} & s_{21} & 0 \end{bmatrix} \tag{6.57}$$

The only remaining independent scattering coefficient s_{21} can be determined with the help of the unitary condition Equation 6.53, applied to any one of the columns of the matrix (6.57).

$$2|\, s_{21}|^2 = 1 \tag{6.58}$$

Equation 6.58 still leaves the phase factor included in s_{21} undetermined. Provided the reference planes of the junction are suitably chosen this can be assumed to be unity. Hence the final form of the scattering matrix is

$$S = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 & e^{-j\pi/4} & 0 \\ 1 & 0 & 0 & e^{-j\pi/4} \\ e^{-j\pi/4} & 0 & 0 & 1 \\ 0 & e^{-j\pi/4} & 1 & 0 \end{bmatrix} \tag{6.59}$$

6.5 SCATTERING MATRICES OF NONRECIPROCAL LINEAR JUNCTIONS

In the preceding section the scattering matrix of a reciprocal linear junction was derived, starting from the impedance matrix. As a result of the symmetry of the impedance matrix it was shown that the scattering matrix of such a junction is also symmetric.

In the discussion of two-port networks in Chapter 3, it was found that impedance matrices of non-reciprocal networks are no longer symmetric. The same observation can be made about the impedance parameters of multiport networks, as discussed in Section 6.1. As soon as active, or otherwise non-reciprocal elements are included in the circuit, the symmetry of the impedance matrix disappears. In view of this fact, Rule 3, Section 6.3, no longer applies. In consequence, the *scattering matrices of non-reciprocal junctions are not symmetric*. This is the main property of the wave matrices of such microwave devices as ferrite isolators or circulators.

Even in the absence of reciprocity the scattering coefficients of a junction remain functions of frequency only, just like the corresponding impedance parameters, as long as the junction is operated under linear conditions. The present discussion is restricted to such cases.

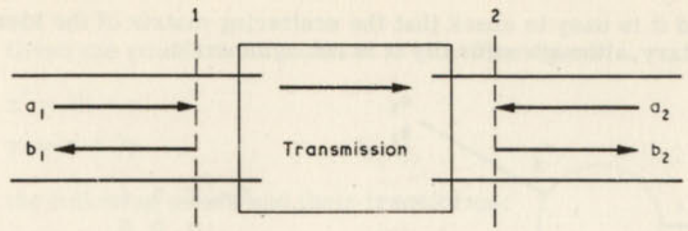
To sum up we can write for a linear, non-reciprocal junction

$$S^t \neq S \tag{6.60}$$

In the preceding section it was also established that the scattering matrices of lossless junctions are unitary. Since the proof of this property in no way depends on the symmetry of the matrix, it must apply to non-reciprocal junctions as well as reciprocal ones.

To exemplify the foregoing statements let us obtain the scattering matrices of an idealised ferrite isolator and a three-port circulator.

Fig. 6.10 shows schematically an idealised isolator, or one way transmission line. All energy entering the input port is assumed to pass through the device without reflection or attenuation. On the other hand the isolator prevents any power to pass to the left, but it does this without reflection. The assumption



$$S = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

Fig. 6.10

of no reflections inside the two-port allows us to deduce that the diagonal elements of its scattering matrix vanish. (see p. 197).

$$s_{11} = s_{22} = 0 \tag{6.61}$$

Furthermore, since a wave of amplitude a_1 applied to the input, appears unchanged at the output and is then denoted by b_2 we find

$$s_{21} = 1 \tag{6.62}$$

provided the reference planes of the ports are so chosen that the phase factor included in s_{21} becomes unity. Finally, as nothing appears at the input in response to a wave of amplitude a_2 applied to the output, we obtain

$$s_{12} = 0 \tag{6.63}$$

The results are summarised in Fig. 6.10. As was to be expected the scattering matrix of the isolator is not symmetric.

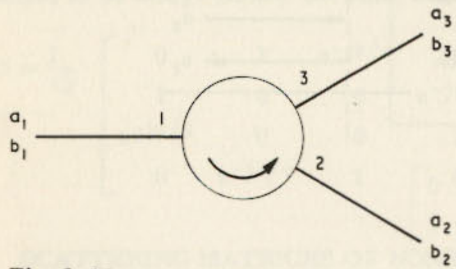
Fig. 6.11 shows schematically an idealised three-port circulator. The ports are indicated diagrammatically by single lines, and the direction of propagation within the device is marked by an arrow. The scattering matrix of this device is of order 3 × 3, and is easily found by inspection, under similar assumptions as were made regarding the isolator. In the first place all diagonal elements are zero. Then we find

$$s_{21} = s_{32} = s_{13} = 1 \tag{6.64}$$

All the remaining scattering coefficients are zero, and the complete matrix is presented in Fig. 6.11.

It is interesting to note that the 2 × 2 matrix of the isolator is not unitary, indicating that a perfect one way transmission line cannot be made lossfree.

On the other hand it is easy to check that the scattering matrix of the ideal circulator is unitary, although naturally it is not symmetric.



$$S = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Fig. 6.11

References on Chapter 6

MONTGOMERY, C. G., DICKE, R. H., PURCEL, E. M.
'Principles of Microwave Circuits', *M. I. T. Radiation Laboratory Series*, Vol. 8, McGraw-Hill.

LAX, B., BUTTON, K. J.
Microwave Ferrites and Ferrimagnetics, McGraw-Hill, (1962).

MATHEWS, E. W.
'The Use of Scattering Matrices in Microwave Circuits', *I.R.E. Trans.* (Microwave Theory and Techniques) (April 1955).

BOLINDER, E. F.
'Note on the Matrix Representation of Linear Two Port Networks', *Trans.I.R.E.* (Circuit Theory) (December 1957).

1. Given the point transformation in two dimensions

$$x' = 2x + y$$
$$y' = x + 3y$$

plot the following points and their transforms:

$$(0,0),\ (1,0),\ (0,1),\ (1,1),\ (0,\infty).$$

2. Carry out the following matrix multiplications:

$$\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & -2 & 3 \\ -4 & 2 & 5 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ -1 & 0 \\ 2 & 4 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix},$$

$$\begin{bmatrix} 1+i & 1-i \end{bmatrix}\begin{bmatrix} 1-i \\ 1+i \end{bmatrix}, \quad \begin{bmatrix} e^{i\psi} & e^{-i\psi} \end{bmatrix}\begin{bmatrix} e^{-i\psi} \\ e^{i\psi} \end{bmatrix},$$

$$\begin{bmatrix} 1 & 0 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} 0 \\ 1/2 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 2 \\ 1/2 \end{bmatrix}\begin{bmatrix} 1 & 0 & 2 \end{bmatrix}$$

3. Two point transformations in the plane are defined by the matrices

$$\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}, \quad \begin{bmatrix} 0 & 2 \\ 1/2 & 1 \end{bmatrix},$$

Apply these transformations to the point $(1,1)$ in turn, and then in reverse order. Plot the results. Satisfy yourself that they agree with the two product transformations obtained on multiplying the matrices together in the two possible sequences.

4. Given the matrices

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

apply the rules for the multiplication of matrices, addition of matrices and multiplication by a scalar factor to establish the following relations:

$$A^2 = B^2 = C^2 = I$$
$$AB + BA = BC + CB = CA + AC = 0$$
$$AB - BA = 2jC, \; BC - CB = 2jA, \; CA - AC = 2jB$$

5. Check the following product of diagonal matrices

$$\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}\begin{bmatrix} 1/a & 0 & 0 \\ 0 & 1/b & 0 \\ 0 & 0 & 1/c \end{bmatrix} = I$$

6. Given a symmetric matrix of order 2×2 and a vector of coordinates in the plane, form the following product:

$$\begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} a & h \\ h & b \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

Note that the result is a scalar expression which, when equated to unity, is the equation of a quadratic curve. Satisfy yourself that the matrix form of the equation of a quadric surface in 3 dimensions is

$$\begin{bmatrix} x & y & z \end{bmatrix}\begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = 1$$

7. Form the adjoint matrix and the reciprocal matrix of the following matrices:

$$\begin{bmatrix} 1 & 0 & 2 \\ 3 & 0 & 0 \\ 1 & 2 & 0 \end{bmatrix}; \quad \begin{bmatrix} 2 & 1 & 3 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \quad \begin{bmatrix} p & 0 & 0 \\ 0 & q & 0 \\ 0 & 0 & r \end{bmatrix}; \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

8. Solve the following sets of linear equations by evaluating the reciprocal matrix of each system

$$x + y = 1; \quad 2x + y = 1; \quad 1/4x - y = 0;$$
$$2x + 1/2y = 2 \quad 3x + 2y = 0 \quad x - y = 3$$

$$2x - z = 0$$
$$3x - 1/2z = 2$$
$$-x + 3y + z = 1$$

9. Given the point transformation

$$x' = x + y$$
$$y' = 2x + 1/2y$$

find the points in the OXY– plane which transform into the following points in the OX′Y′– plane

$$(0,1), \; (1,0), \; (1,1), \; (1,2), \; (1, 1/2).$$

10. Expand $(A + B + C)^2$ into a sum of squares and compare with the corresponding scalar expression.

11. Given the matrix

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}$$

check that it satisfies the relation

$$A^2 - A - 6I = O$$

12. The transformation of a locus of points in the plane can be accomplished as follows. Given the equation of the locus in explicit form $y = f(x)$, substitute for y on the right side of the equations of transformation.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x \\ f(x) \end{bmatrix}$$

The equation of the transformed locus in the OX′Y′– plane is now given in parametric form, with x as the parameter. On elimination of x the equation assumes the usual form $y' = F(x')$

(*a*) Apply this procedure to the parabola $y = x^2$ under the transformation of Example 9.
Repeat for the transformation

$$x' = y, \quad y' = -x.$$

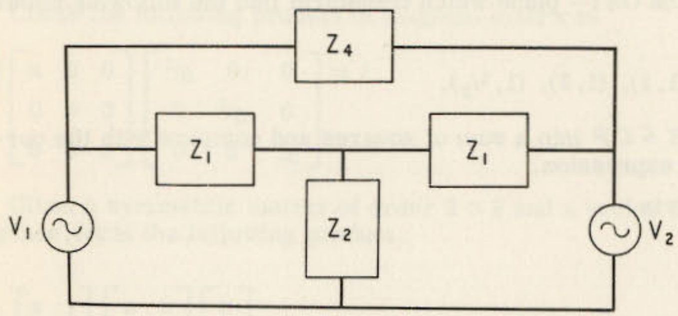
(*b*) Given the straight line $x' + 2y' + 1 = 0$ in the OX′Y′–plane find the locus corresponding to it in the OXY–plane under the transformation of Example 9. Do the same for the transformation

$$x' = \frac{\sqrt{3}}{2}x + \frac{1}{2}y$$
$$y' = -\frac{1}{2}x + \frac{\sqrt{3}}{2}y$$

(*c*) Repeat (*b*) for the unit circle $x'^2 + y'^2 = 1$.

EXAMPLES ON CHAPTER 3

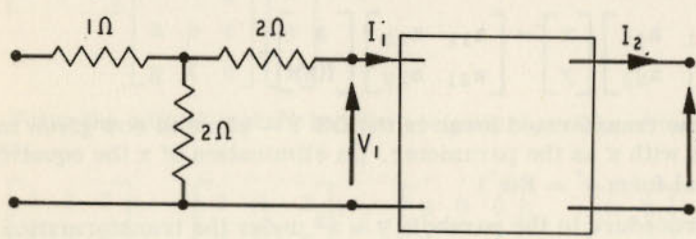
1. Write down in matrix form the mesh equations of the circuit shown. Treating the circuit as a four-terminal network with ports at V_1 and V_2,



A 3.1

solve for the currents in the ports and hence find the y-parameters. Also obtain the z-parameters of the two-port and compare results with Section 26.

2. Two four-terminal networks are connected in cascade as shown in the diagram. In addition to the information marked in the figure the following is

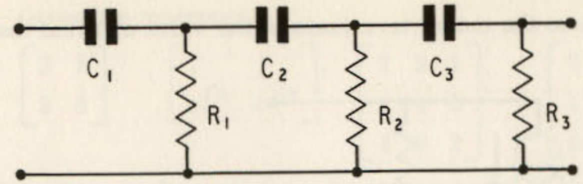


A 3.2

known about the network on right hand side. With port 2 open circuited $\dfrac{V_1}{V_2'} = \dfrac{V_2'}{I_1} = 2$; with port 2 short circuited $\dfrac{I_1}{I_2} = 2$.

(*a*) What are the general circuit constants of the 'black box'?
(*b*) Evaluate the general circuit constants of the cascade connection.
(*c*) Find the equivalent T—circuit of the 'black box'.
(*d*) Obtain the equivalent T— and π— circuits of the combination.

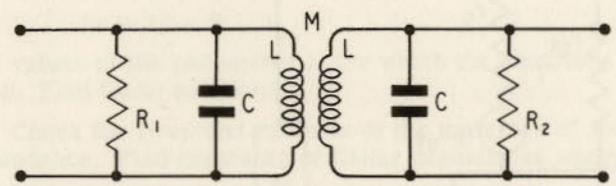
3. The network shown in the diagram is used to feed the anode voltage back to the grid of a vacuum valve phase shift oscillator. Since oscillations are



A 3.3

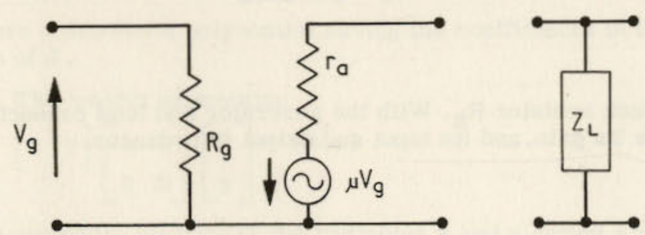
excited only when the terminal voltages are in antiphase, find values for the circuit components to secure this condition.

4. Given the current entering the two-port network shown, find the voltage across its output terminals.



A 3.4

5. The constant voltage source equivalent circuit of a vacuum valve at low frequencies is shown in the accompanying diagram. Write down its matrices of z—parameters and general circuit constants. Connect the load to the out-

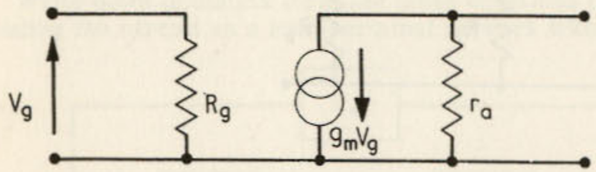


A 3.5

put terminals and find the general circuit constants of the combination. What is the gain of the circuit?

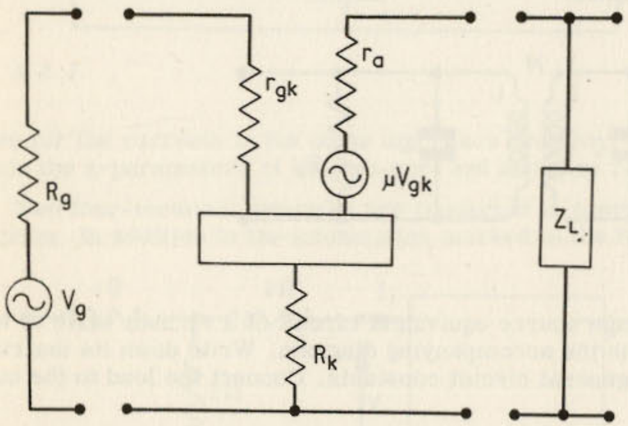
6. The constant current source equivalent circuit of a vacuum valve at low frequencies is given. Write down its y—parameters. Connect the double

tuned circuit of Example 4. to the output terminals of the valve and find the gain at resonance.



A 3.6

7. The figure gives an example of a feedback circuit — the cathode follower. Find the z—parameters of the combined two-port network consisting



A 3.7

of the valve and feedback resistor R_k. With the generator and load connected to the circuit, evaluate its gain, and its input and output impedances.

1. Given the matrices

$$\begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}, \quad \begin{bmatrix} 1 & j \\ e^{j\pi/2} & e^{j\pi/2} \end{bmatrix}, \quad \begin{bmatrix} 2 & 3 & 1 \\ 1 & 4 & 5 \\ 4 & 6 & 2 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

reduce them to diagonal form by a succession of elementary operations; (*a*) on rows only; (*b*) on columns only; (*c*) on both rows and columns. Satisfy yourself that the number of nonvanishing diagonal elements is the same regardless of the method of reduction.

2. Ascertain the rank of the matrices of Example 1 by direct evaluation of their determinants and minors. Check in each case that this equals the number of nonvanishing diagonal elements in the reduced forms found above.

3. Given the system of homogeneous linear equations

$$(1-\lambda)x + 3y = 0$$
$$2x - \lambda y = 0$$

find values of the parameter λ, for which the equations have non-trivial solutions. Find these solutions.

4. Check the rows and columns of the matrices of Example 1. for linear dependence. Find constants of linear dependence where applicable.

5. Verify that the solutions of the system of Example 3., belonging to different values of λ, are linearly independent.

6. Apply the decomposition theorem for functions of matrices to prove the Cayley-Hamilton theorem. This states that a square matrix satisfies its own characteristic (or eigenvalue) equation. In symbols show that

$$\varphi(A) = 0$$

where φ denotes a polynomial having the coefficients of the eigenvalue equation of A.

7. The matrix expression

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & h \\ h & b \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = f(x, y)$$

is a quadratic function of the variables x and y called a quadratic form. When the matrix of a quadratic form is diagonal the form becomes a sum of squares. Apply the methods of Section 4.8 to reduce the above form to a sum of squares.

When equated to unity the above quadratic form becomes the equation of a quadratic curve. Assuming the form is an ellipse find the relation between its semi-axes and the eigenvalues of the matrix of the form.

Suggestions for further reading

AITKEN, A. C.
Determinants and Matrices, Oliver and Boyd, (1964).

GUILLEMIN, E. A.
The Mathematics of Circuit Analysis, John Wiley & Sons. Inc., (1949).

GOERTZEL, G., and TRALLI, N.
Some Mathematical Methods of Physics, McGraw-Hill, (1960).

LANCZOS, C.
Linear Differential Operators, D. Van Nostrand, (1961).

Index

matrix algebra for electronic engineers

Hlawiczka

Iliffe